The Astrophysical Journal, 917:64 (24pp), 2021 August 20 https://doi.org/10.3847/1538-4357/ac088a

The AGORA High-resolution Galaxy Simulations Comparison Project. III. Cosmological Zoom-in Simulation of a Milky Way–mass Halo

Santi Roca-Fàbrega[1,28], Ji-hoon Kim[2,3,28], Loic Hausammann[4,28], Kentaro Nagamine[5,6,7,28], Alessandro Lupi[8,28], Johnny W. Powell[9,28], Ikkoh Shimizu[10,28], Daniel Ceverino[11,12], Joel R. Primack[13], Thomas R. Quinn[14], Yves Revaz[4], Héctor Velázquez[15], Tom Abel[16,17,18], Michael Buehlmann[19], Avishai Dekel[20], Bili Dong[21], Oliver Hahn[19], Cameron Hummels[22], Ki-won Kim[2], Britton D. Smith[23], Clayton Strawn[13], Romain Teyssier[24], and Matthew J. Turk[25]
The AGORA Collaboration[26,27]

[1] Departamento de Física de la Tierra y Astrofísica, Facultad de Ciencias Físicas, Plaza Ciencias, 1, E-28040 Madrid, Spain; sroca01@ucm.es
[2] Center for Theoretical Physics, Department of Physics and Astronomy, Seoul National University, Seoul 08826, Republic of Korea; mornkr@snu.ac.kr
[3] Seoul National University Astronomy Research Center, Seoul 08826, Republic of Korea
[4] Institute of Physics, Laboratoire d'Astrophysique, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland; loic.hausammann@epfl.ch
[5] Department of Earth and Space Science, Graduate School of Science, Osaka University, Toyonaka, Osaka, 560-0043, Japan; kn@astro-osaka.jp
[6] Kavli IPMU (WPI), University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba, 277-8583, Japan
[7] Department of Physics & Astronomy, University of Nevada Las Vegas, Las Vegas, NV 89154, USA
[8] Dipartimento di Fisica "G. Occhialini," Università degli Studi di Milano-Bicocca, I-20126 Milano, Italy
[9] Department of Physics, Reed College, Portland, OR 97202, USA
[10] Shikoku Gakuin University, 3-2-1 Bunkyocho, Zentsuji, Kagawa, 765-8505, Japan
[11] Departamento de Física Teórica, Facultad de Ciencias, Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain
[12] CIAFF, Facultad de Ciencias, Universidad Autónoma de Madrid, E-28049 Madrid, Spain
[13] Department of Physics, University of California at Santa Cruz, Santa Cruz, CA 95064, USA
[14] Department of Astronomy, University of Washington, Seattle, WA 98195, USA
[15] Instituto de Astronomía, Universidad Nacional Autónoma de México, A.P. 70-264, 04510, Mexico, D.F., Mexico
[16] Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA 94305, USA
[17] Department of Physics, Stanford University, Stanford, CA 94305, USA
[18] SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
[19] Laboratoire Lagrange, Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Boulevard de l'Observatoire, CS 34229, F-06304 Nice, France
[20] Center for Astrophysics and Planetary Science, Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel
[21] Department of Physics, Center for Astrophysics and Space Sciences, University of California at San Diego, La Jolla, CA 92093, USA
[22] TAPIR, Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
[23] Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK
[24] Centre for Theoretical Astrophysics and Cosmology, Institute for Computational Science, University of Zurich, Zurich, 8057, Switzerland
[25] School of Information Sciences and Department of Astronomy, University of Illinois, Urbana-Champaign, IL 61820, USA
Received 2021 March 30; revised 2021 June 2; accepted 2021 June 4; published 2021 August 17

Abstract

We present a suite of high-resolution cosmological zoom-in simulations to $z = 4$ of a $10^{12}\,M_{\odot}$ halo at $z = 0$, obtained using seven contemporary astrophysical simulation codes (ART-I, ENZO, RAMSES, CHANGA, GADGET-3, GEAR, and GIZMO) widely used in the numerical galaxy formation community. The physics prescriptions for gas cooling and heating and star formation are the same as the ones used in our previous Assembling Galaxies of Resolved Anatomy (AGORA) disk comparison but now account for the effects of cosmological processes such as the expansion of the universe, intergalactic gas inflow, and the cosmic ultraviolet background radiation emitted by massive stars and quasars. In this work, we introduce the most careful comparison yet of galaxy formation simulations run by different code groups, together with a series of four calibration steps each of which is designed to reduce the number of tunable simulation parameters adopted in the final run. In the first two steps, we methodically calibrate the gas physics, such as cooling and heating, in simulations without star formation. In the third step, we seek agreement on the total stellar mass produced with the common star formation prescription used in the AGORA disk comparison, in stellar-feedback-free simulations. In the last calibration step, we activate stellar feedback, where each code group is asked to set the feedback prescription to as close to the most widely used one in its code community as possible, while aiming for convergence in the stellar mass at $z = 4$ to the values predicted by semiempirical models. After all the participating code groups successfully complete the calibration steps, we achieve a suite of cosmological simulations with similar mass assembly histories down to $z = 4$. With numerical accuracy that resolves the internal structure of a target halo ($\lesssim$100 physical pc at $z = 4$), we find that the codes overall agree well with one another, e.g., in gas and stellar properties, but also show differences, e.g., in circumgalactic medium (CGM) properties. We argue that, if adequately tested in accordance with our proposed calibration steps and common parameters, high-resolution cosmological zoom-in simulations can have robust and reproducible results. New code groups are invited to join and enrich this comparison by generating equivalent models or to test the code's compatibility on their own, by adopting the common initial conditions, the common easy-to-implement physics package, and the proposed calibration steps. Further analyses of the zoom-in

[26] http://www.AGORAsimulations.org
[27] The authors marked with "28" as code leaders contributed to the article by leading the effort within each code group to perform and analyze simulations.
[28] Code leaders.

simulations presented here will be presented in forthcoming reports from the AGORA Collaboration, including studies of the CGM, simulations by additional codes, and results at lower redshift.

Unified Astronomy Thesaurus concepts: Galaxy formation (595); Astronomical simulations (1857); Hydrodynamical simulations (767); Stellar feedback (1602)

1. Introduction

Established in 2012, the Assembling Galaxies of Resolved Anatomy (AGORA) High-resolution Galaxy Simulations Comparison Project has aimed at collectively raising the predictive power of contemporary numerical galaxy formation studies, by carefully comparing high-resolution galaxy simulations on multiple code platforms widely used in the field. The main goal of the AGORA initiative has been to ensure that physical assumptions are responsible for any success in numerical studies, rather than manifestations of a particular numerical implementation. As of this writing, we have more than 160 individuals from over 60 different academic institutions worldwide who have agreed to the project's philosophy and participated in its collaborative effort in varying degrees. The collaboration has provided a sustainable platform on which members could talk to and learn from others from different code communities, and discuss ambitious "multiplatform" collaborations. The project indeed has become a great social experiment in itself—about the scientific community's collective willingness to ensure the integrity and reproducibility of its experiments.[29]

The first paper of the collaboration (Kim et al. 2014, hereafter Paper I) focused on introducing the project to the community. It presented the first proof-of-concept simulations, dark-matter-only but using cosmological zoom-in initial conditions. Results from comparing cosmological simulations among nine flavors of state-of-the-art numerical codes showed a robust convergence. In a second paper from the AGORA Collaboration (Kim et al. 2016, hereafter Paper II) we presented a comparison of idealized Milky Way–mass galaxies simulated in isolation, obtained from nine widely used state-of-the-art gravitohydrodynamics codes, which have recently been made available to be freely used by the community (Roca-Fàbrega et al. 2020). The simulations in Paper II achieved overall agreement with one another in many parameter spaces for both gaseous and stellar components. Yet, some discrepancies were expected and present, which were understood as systematic differences between codes—for example, between mesh-based and particle-based codes in low-density regions, and between more diffusive and less diffusive schemes in high-density regions. Such intrinsic differences were, however, found to be small in general compared to variations in implementations of common subgrid physics, such as supernova (SN) feedback.

The AGORA Project has helped to establish a simulation infrastructure essential to our achieving thorough comparisons thus far, and it will allow and foster future comparisons. This infrastructure includes, among others, a common initial condition generator (MUSIC; Hahn & Abel 2011),[30] a common gas cooling and heating scheme (GRACKLE; Smith et al. 2017),[31] and a common analysis toolkit (`yt`; Turk et al. 2011),[32] all of which are publicly available software. In particular, all the figures and plots in this article and Papers I and II have been produced with the AGORA common analysis platform based on `yt`. It is also worth noting that several recent comparison and calibration studies have been motivated by the results presented in our previous reports. Examples include a study of changes in star formation efficiency in molecular clouds (Grisdale et al. 2019) and tests of new star formation and SN feedback implementations, in both isolated (Shimizu et al. 2019) and cosmological contexts (Oh et al. 2020).

Building on these past achievements, in this third paper of our continuing endeavor in AGORA, we follow a path similar to that of Paper II, but this time with cosmological zoom-in simulations. This type of comparison has never been properly carried out due to its complexity and time-consuming nature. However, it is now possible—though still challenging—thanks to the infrastructure the AGORA Collaboration has built and maintained. A reproducibility check like this is essential as the field relies increasingly on the numerical verification of galaxy formation theories in cosmological contexts. All code groups start their simulations from a common initial condition generated with MUSIC (Section 2). The physics prescriptions (e.g., gas cooling and heating and star formation parameters) are also common to all participating codes as in Paper II, although some changes are made in each code (Section 3). Only the decision concerning the stellar feedback prescription and metal production to be used is left to each code group, and the code groups are asked to use a prescription close to the most widely used practice in each code community. Spatial resolution of $\lesssim$100 physical pc at $z = 4$ is imposed to resolve the internal structure of a target halo, and to make our physics prescriptions less reliant on platform-specific models (Section 4). After a series of calibration steps for the adopted physical processes (Figure 1 and Section 5), we reach a suite of simulations illustrating how seven state-of-the-art codes reproduce the formation and evolution of a Milky Way–type galaxy in a cosmological context down to $z = 4$ with their favorite stellar feedback and metal production prescriptions (Section 6). As in previous AGORA comparisons, we caution that we do not intend to identify a correct or incorrect code, but to focus on juxtaposing different codes for physical insights and learn how much scatter one should expect among modern simulations.

This paper is organized as follows. Section 2 describes the initial conditions of our experiment. We discuss the physics modules employed in our simulations in Section 3, and the runtime parameters in Section 4. Section 5 presents our calibration steps designed to prepare the ground for the final simulation entries. In Section 6 we compare the results of our final runs, focusing on the stellar and gas properties of the target halo, and its evolution in time. Finally, in Section 7 we conclude the article with remarks on how AGORA's multiplatform approach can significantly enhance the scientific value of numerical galaxy formation studies.

[29] See the project website at http://www.AGORAsimulations.org/ for more information about the AGORA Collaboration.
[30] The website is https://www-n.oca.eu/ohahn/MUSIC/.
[31] The website is http://grackle.readthedocs.io/.
[32] The website is http://yt-project.org/.

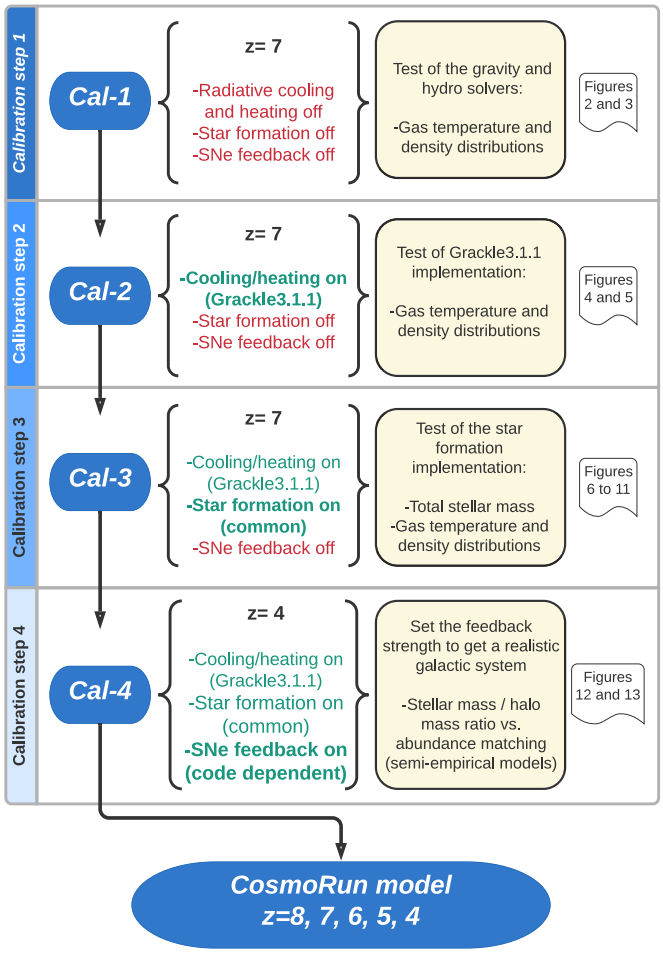

Figure 1. Summary of the physics calibration procedure. We indicate, from left to the right, the target redshift and the physics prescriptions in each step, the main objective and the variables used to test convergence, and the corresponding figures.

2. Initial Conditions

We use a set of parameters for MUSIC, an initial condition generator with an adaptive multigrid Poisson solver (Hahn & Abel 2011), that depicts a halo evolving to a virial mass of $\sim 10^{12}\,M_{\odot}$ at $z = 0$ with a relatively quiescent merger history between $z = 2$ and 0.[33] The initial condition, tagged 1e12q, is identified and made publicly available by the AGORA Collaboration (Paper I).[34] We assume a flat ΛCDM cosmology consistent with WMAP7/9+SNe+BAO: $\Omega_{\rm m} = 0.272$, $\Omega_{\Lambda} = 0.728$, $\sigma_8 = 0.807$, $n_{\rm s} = 0.961$, and $H_0 = 70.2\ {\rm km\ s^{-1}\ Mpc^{-1}}$ (Komatsu et al. 2011; Hinshaw et al. 2013). The initial metallicity is set to $10^{-4}\,Z_{\odot}$ everywhere.[35]

With a 128^3 root resolution in a $(60\ \text{comoving}\ h^{-1}\ \text{Mpc})^3$ box and a series of five nested higher-resolution regions, the equivalent unigrid resolution at the finest zoom-in region is 4096^3 (i.e., MUSIC parameters $[\ell_{\rm min}, \ell_{\rm max}] = [7, 12]$). The highest-resolution region in this initial condition is in an ellipsoidal shape that is large enough to enclose all the particles that eventually end up within $4R_{\rm vir}$ of the target halo at $z = 0$. Correspondingly, the target halo contains the highest-resolution particles of masses $m_{\rm DM,IC} = 2.8 \times 10^5\,M_{\odot}$ and $m_{\rm gas,IC} = 5.65 \times 10^4\,M_{\odot}$, with the latter designed to approximately match the gas resolution in Paper II, $m_{\rm gas} = 8.6 \times 10^4\,M_{\odot}$. For more information about this initial condition and other available AGORA initial conditions, we refer the interested reader to Section 2 of Paper I.

[33] Here we use MUSIC's changeset ID `eb870ed`.

[34] See http://www.AGORAsimulations.org/ or http://sites.google.com/site/santacruzcomparisonproject/blogs/quicklinks/.

[35] $1\,Z_{\odot} = 0.02041$ is used across all participating codes in order to follow our choice in Paper II (see Section 2 of Paper II for details).

3. Physics in the Codes

We briefly summarize the key physics and code-by-code differences for this particular comparison.

3.1. Common, Code-independent Physics

The common baryonic physics for our study is based on Papers I and II. First, the cooling library GRACKLE determines the rate of radiative gas cooling based on the properties of the gas parcels (Smith et al. 2017). The interface we built for Paper II is utilized by each participating code, in the equilibrium cooling mode of GRACKLE V3.1.1. Here, GRACKLE looks up a precomputed CLOUDY cooling table for primordial and $1\,Z_{\odot}$ metallicities as functions of gas density and temperature (Ferland et al. 2013). To obtain the corresponding gas cooling and heating rates, the $1\,Z_{\odot}$ rates are linearly scaled by the gas metallicity (Section 3.1 of Paper II), and the result is added to the values from the primordial gas to get the combined rate. GRACKLE also includes redshift-dependent cosmic UV background (UVB) radiation (Haardt & Madau 2012) with hydrogen self-shielding (i.e., input file `CloudyData_UVB=HM2012_shielded.h5`; see also Section 3.3 of Paper I). In addition, instead of using GRACKLE's own cosmic microwave background (CMB) temperature floor, each code is supplemented with a redshift-dependent but density-independent CMB floor.[36]

Lastly, in order to prevent unphysical collapse or fragmentation due to limited resolution, in calibration steps 3 and 4 we apply a nonthermal pressure floor $P_{\rm Jeans}$ that forces the local Jeans length to be resolved at a given numerical resolution at all times. Its value is $P_{\rm Jeans} = (\gamma\pi)^{-1} N_{\rm Jeans}^2 G \rho_{\rm gas}^2 \Delta x^2$, where $\gamma = 5/3$ is the adiabatic index, $N_{\rm Jeans} = 4$ is the Jeans number, G is the gravitational constant, $\rho_{\rm gas}$ is the gas density, and Δx is the finest spatial resolution in physical units (the finest cell size for mesh-based codes, or the gravitational softening length for particle-based codes; see Section 4). This additional pressure term can be interpreted as an extra pressure source due to unresolved interstellar medium (ISM) turbulence. For actual implementations of the pressure floor in each code, we refer the reader to Section 3.1 and Appendix A of Paper II.

When the density of a gas parcel exceeds $n_{\rm H,thres} = 1\ {\rm cm^{-3}}$ (note the difference from the $n_{\rm H,thres}$ used in Paper II), a star particle can be created at a rate of $d\rho_{\star}/dt = \epsilon_{\star}\rho_{\rm gas}/t_{\rm ff}$, where $\epsilon_{\star} = 0.01$ is the formation efficiency and $t_{\rm ff} = (3\pi/(32 G \rho_{\rm gas}))^{1/2}$ is the local freefall time. The only freedom that is left to each code group is to choose the stochastic or deterministic nature of this process. A single star particle depicts a collection of cluster-sized masses sharing the same age and metallicity, corresponding to a single stellar population. It is required to weigh more than $6.1 \times 10^4\,M_{\odot}$ at creation for mesh-based codes—a value approximately matching the gas resolution in the initial condition, $m_{\rm gas,IC} = 5.65 \times 10^4\,M_{\odot}$—or to inherit the mass of its parent gas particle in particle-based codes. In Paper II, our stellar feedback formula implied one Type II SN event per 91 $M_{\odot}$

[36] This functionality is planned to be added to the latest GRACKLE.

Table 1
Stellar Feedback Implementation Adopted by Each Code Group[a]

Code	Stellar Feedback	SN and Metal Production Model	Effective Metal Yield	Runtime Parameters
ART-I	T + K, RP	SNe Ia/II, AGB stars*	0.033	$E_{\rm thermal} = 2 \times 10^{51}$ erg/SN, $p = 3.6 \times 10^6\ M_\odot$ km s^{-1}/SN
ENZO	T	SNe II	0.032	$E_{\rm thermal} = 5 \times 10^{52}$ erg/SN
RAMSES	T, DC	SNe II	0.033	$E_{\rm thermal} = 4 \times 10^{51}$ erg/SN, $\sigma_{\rm min} = 100$ km s^{-1}, $T_{\rm delay} = 10$ Myr
CHANGA	T + S	SNe Ia/II, AGB stars**	0.032	$E_{\rm thermal} = 5 \times 10^{51}$ erg/SN
GADGET-3	T + K, RP, DC	SNe Ia/II, AGB stars	0.025	$E_{\rm SN} = 4 \times 10^{49}$ erg/$M_\odot$, $T_{\rm delay} = t_{\rm hot}$ (see Section 3.2.5)
GEAR	T, DC	SNe Ia/II	0.024	$E_{\rm thermal} = 4.5 \times 10^{51}$ erg/SN, $T_{\rm delay} = 5$ Myr
GIZMO	T + K	SNe II	0.033	$E_{\rm SN} = 5 \times 10^{51}$ erg/SN

Note.
[a] T = thermal feedback, K = kinetic feedback, RP = radiation pressure, DC = delayed cooling, S = superbubble, * = only for energy production (not metal), and ** = only for metal production (not energy). While the total returned mass via feedback is constrained across the code platforms (Section 3.1), the exact feedback scheme and the metal yield are left to each code group's discretion to be as close to the most widely used practice in its community as possible. For more information on the items listed here, see Section 3.2. For more information on the effective metal yield by stellar feedback measured in the entire simulation box at $z = 4$ for the `CosmoRun` suite of simulations (fourth column), see Section 6.2.2.

stellar mass formed, each instantaneously releasing 10^{51} erg of thermal energy, 14.8 $M_\odot$ of gas, and 2.6 $M_\odot$ of metals. In contrast, in this work, while the returned mass is equal to that in Paper II, the exact deposit scheme into the ISM, such as stellar winds or SN events, and its associated energy and metal yields are left to each code group's discretion. We do ask the deposit scheme be as close to the most widely used practice in the community as possible (detailed in Section 3.2 and Table 1; see also Sections 3.2 and 5 and Appendix B of Paper II). We also leave the choice of implementing an explicit metal diffusion scheme to each particle-based code group (see Sections 3.2.4 to 3.2.6 for details).

We note that our common physics models, including the subgrid physics (e.g., star formation), helped us in Paper II to produce similar stellar disks across all codes—comparable in terms of their morphologies, kinematics, and star formation relations, to name a few (Sections 6.4–6.6 of Paper II). In the present comparison, however, we use a fully cosmological setup that is substantially more complex. Although the common subgrid physics models here are based on the ones in the idealized galaxy setup (Paper II), we have found a need to introduce changes to the fiducial parameters to reproduce a realistic galactic system at low redshift. The fiducial set of parameters has been modified thus: the star formation threshold density $n_{\rm H,thres} = 1\ {\rm cm}^{-3}$ instead of the $10\ {\rm cm}^{-3}$ in Paper II and a stellar feedback scheme instead of the common simple thermal deposit model in Paper II (see Section 3.2). These changes have been motivated by the deviation of $M_*/M_{\rm halo}$ from the observed value in Paper II, and by a need to account for the potential redshift dependence of the adopted physics.

3.2. Participating Codes and Code-dependent Physics

Here we briefly explain the physics included in each code, focusing only on the part that is changed from Paper II, or is unique for each code. Hence, interested readers are encouraged to see our previous work to grasp the full picture of how each code works—Paper I for gravitational dynamics and Paper II for hydrodynamics. In particular, Table 1 summarizes the key stellar feedback parameters and effective metal yield in each code, for which each code group is given freedom to choose its own feedback scheme for energy and metals. It should be noted that the code groups to be involved in future AGORA studies are not limited to the seven codes listed in this section.

3.2.1. ART-I

The ART-I code (Kravtsov et al. 1997; Kravtsov 2003; Ceverino & Klypin 2009) used to obtain the cosmological simulation presented here is based on the one used in the previous comparison efforts (Papers I and II). Only a few minor modifications should be noted. Among them is a change in the adaptive mesh refinement (AMR) strategy to better follow the cosmic evolution of large-scale structures. This change is in line with what has been commonly used in previous ART-I cosmological zoom-in simulations (e.g., Ceverino et al. 2010, 2014, 2017). We have also updated the gas cooling and heating scheme from ART-I's own implementation using the CLOUDY table in Paper II to the standard-package GRACKLE V3.1.1 in the current paper. The nonthermal pressure floor in ART-I is slightly different from the common prescription (Section 3.1); in other words, the Jeans length is resolved by at least seven resolution elements at all times (Ceverino et al. 2010).

ART-I uses a stochastic star formation subgrid model. Details on this star formation model can be found in Ceverino et al. (2014). We slightly change the stochasticity of star formation to ensure that we use the common star formation efficiency value (Section 3.1). ART-I's prescription fits within the agreed AGORA parameter range. The treatment of stellar feedback is similar to the model in Ceverino et al. (2017), which includes thermal, kinetic, and radiation pressure feedback. The code also includes the later effects of SNe Ia and stellar mass loss, and it follows the metal enrichment of the ISM. The convergence goal in calibration step 4 (`Cal-4`; Section 5.4) is achieved by the widely used feedback model in the `VELA6` simulations (D. Ceverino et al., in preparation) but with four times more injection of momentum (see parameter p in Table 1). This increase compensates for the differences in resolution. The default AGORA effective metal yield has been obtained by increasing the standard SN II and SN Ia yields in ART-I by a factor of four.

3.2.2. ENZO

The ENZO code (Bryan et al. 2014; Brummel-Smith et al. 2019) for this work is from the `master` branch in the publicly available `enzo-dev` repository.[37] Star formation is implemented following the same approach in Paper II, which is a

[37] The website is http://enzo-project.org/. Here we use ENZO's changeset ID `02c88172`.

fully deterministic scheme. To incorporate the stellar feedback model established in Paper II, files such as `star_maker4.F` and `Grid_StarParticleHandler.C` in the said repository need a minor modification. To reach the convergence in our calibration step 4 (`Cal-4`; Section 5.4), the stellar feedback efficiency parameter is increased from the value in Paper II, matching the findings in recent ENZO calibration studies against observations (e.g., Oh et al. 2020; see also Table 1). The model only accounts for effects by SNe II. Other adopted schemes, such as the hydrodynamics solver, are the same as those in Paper II, and are largely in line with recent numerical galaxy formation studies using ENZO (e.g., Kim et al. 2019; Shin et al. 2020).

In order to realize the ellipsoid-shaped initial condition in simulations (Section 2), ENZO identifies and tracks the ellipsoidal Lagrangian region using a special type of dark matter particle called `MustRefineParticle`, which eventually constitutes the target halo at a predetermined target redshift. Cells around these particles are always refined at least down to 20.9 comoving kpc—or five additional refinement levels for a 128^3 root resolution in a (60 comoving h^{-1} Mpc)^3 box—corresponding to the MUSIC parameter $\ell_{\max} = 12$.

3.2.3. RAMSES

The RAMSES code (Teyssier 2002) used in this comparison is from the 2019 December `master` branch of the code repository.[38] Star formation is implemented following Paper II, but without using a temperature threshold. This temperature threshold was closely linked with the implementation of a temperature polytrope to avoid numerical fragmentation, and this approach is no longer used in the present work. Thus, the implementation of nonthermal pressure support to avoid artificial fragmentation takes a different approach from the one in Paper II, being now consistent with the common implementation presented in Section 3.1. With this implementation we ensure that the local Jeans length is resolved at least by four AMR cells at all times. The star formation approach is well described in the most recent works within the code community (e.g., Nuñez-Castiñeyra et al. 2021).

The treatment of stellar feedback here closely follows the so-called "delayed cooling thermal feedback model" formulated in Dubois et al. (2015), and only accounts for effects by SNe II. The RAMSES simulation presented here includes modifications to the model, however, as described in Rosdahl et al. (2017, Section 3.3) and Nuñez-Castiñeyra et al. (2021, Section 2.1.3). Our choices of runtime parameters are listed in Table 1. We note that, out of our tested feedback prescriptions available in RAMSES, the one used here is what succeeded in producing the target stellar mass at $z = 4$ in our calibration step 4 (`Cal-4`; Section 5.4).

3.2.4. CHANGA

CHANGA V3.4 is a reimplementation of the smoothed particle hydrodynamics (SPH) code GASOLINE (Wadsley et al. 2017) in the CHARM++ (CHARMPPPWCPP96) v6.9 runtime system.[39] The code used for the present paper is based on the one in the previous hydrodynamic comparison; therefore, we refer the interested reader to Section 5.5 of Paper II and here we note only a few points and changes. In CHANGA, the kth nearest neighbor algorithm is used to find the $N_{\rm ngb}$th ($= 64$th) nearest neighbors, and then the Wendland C4 kernel (Dehnen & Aly 2012) is employed to determine the hydrodynamic properties. Energy and metals are diffused using the scheme of Shen et al. (2010). We implement GRACKLE V3.1.1 after careful scrutiny.[40]

The treatment of stellar feedback follows the "superbubble" strategy presented by Keller et al. (2014), in contrast to that in Paper II. It includes thermal conduction inside resolved hot bubbles, which maintains uniform temperatures (see the characteristic bubble shapes in Figure 16). This method makes the amount of cold gas heated by feedback not a free parameter, but one set by the thermal conduction. In the first few megayears of feedback heating, the mass contained within a hot bubble can be smaller than the simulation's gas mass resolution, which could result in strong overcooling. To prevent overcooling, the resolution elements briefly represent two components: (1) a hot interior (bubble) where the feedback energy is injected, and (2) a cold shell in pressure equilibrium with the hot interior. The particle returns to a single phase once all the cold gas is evaporated or the hot phase cools below 10^5 K. Thermal energy representing SNe Ia and II is deposited to the neighboring $N_{\rm ngb}$ particles. SN II rates are calculated from the Raiteri et al. (1996) fit to the Padova stellar models. Type Ia rates are computed from the evolution timescales of secondaries in binaries (Matteucci & Greggio 1986). To reach the convergence in our calibration step 4 (`Cal-4`; Section 5.4), the thermal energy is increased to 5×10^{51} erg/SN for the Kroupa initial mass function (IMF), from the typical value used in the community. Metals are released by SNe and asymptotic giant branch (AGB) stars following Raiteri et al. (1996).

3.2.5. GADGET-3

GADGET-3-OSAKA is a modified version of GADGET-3—which itself is an extended version of the SPH code GADGET-2 (Springel 2005). The code includes the common cooling and star formation model detailed in Papers I and II, and the treatment of stellar feedback presented in Aoyama et al. (2017, 2018) and Shimizu et al. (2019). It also includes important improvements, such as the density-independent, pressure–entropy formulation of SPH (Hopkins 2013; Saitoh & Makino 2013), the timestep limiter (Saitoh & Makino 2009), and the quintic spline kernel (Morris 1996), and the number of neighbor particles for each SPH particle is set to 128 ± 8.

For stellar feedback, we distribute both thermal and kinetic energy to neighboring gas particles within a hot bubble, whose size is determined by the local gas density, ambient gas pressure, and feedback energy (see Equations (6)–(7) in Shimizu et al. 2019). We utilize the CELIB chemical evolution library (Saitoh 2017), which provides the chemical yield distribution as a function of time for a given IMF. We deposit metals and energy according to the CELIB output with certain time delays ($t_{\rm hot}$) that depend on the feedback energy, density, and ambient gas pressure, treating SN Ia, Sn II, and AGB star contributions separately.[41] The total injected energy is slightly

[38] The website is https://bitbucket.org/rteyssie/ramses/.

[39] The websites are http://github.com/N-BodyShop/changa/ and http://charm.cs.uiuc.edu/.

[40] The CLOUDY table used in CHANGA differs slightly from the one in the other codes, containing the latest update by the GRACKLE developers. This update only affects an unlikely case of very dense gas at very high redshifts, so it does not change the conclusion of the present article.

[41] For example, oxygen production is always dominated by Type II SNe, carbon is dominated by AGB stars after a few hundred megayears, and iron is dominated by Type Ia SNe after 10^8 yr.

boosted over the canonical CELIB output, to 4×10^{49} erg/$M_{\odot}$ of star-forming gas, corresponding to $E_{\rm SN} = 4 \times 10^{51}$ erg/SN for the Chabrier IMF adopted in CELIB. For details, see Shimizu et al. (2019). The exact prescription used in this paper is similar to the fiducial model `K30T70` therein, except for the equal division of SN energy into thermal (50%) and kinetic (50%) components to reach the target stellar mass (`Cal-4`; Section 5.4). Early stellar feedback is also adopted in the form of thermal energy injected before the first SN explodes. Metal diffusion is not implemented as an explicit process, but metals are smoothed over the SPH kernel when computing the metallicity or cooling rate of each gas particle, mimicking the effect of metal diffusion (Okamoto et al. 2005; Tornatore et al. 2007; Wiersma et al. 2009).

3.2.6. GEAR

The GEAR code is a chemodynamical tree SPH code based on GADGET-2 (Springel 2005). Its original version was described in Revaz & Jablonka (2012) with some improvements discussed in Revaz et al. (2016) and Revaz & Jablonka (2018). For the difference between GEAR and the public version of GADGET-2, we refer the interested reader to Section 5.8 of Paper II. The cooling and star formation prescriptions adopted here are similar to the ones in Paper II.

In our feedback prescription, both energy and yields are deposited among the nearest gas particles so that each neighbor receives a fraction of energy weighted by the SPH kernel. $N_{\rm ngb}$ corresponds to a weighted number of neighbors and is set to 50. Thus, depending on the spatial distribution of gas particles more or less 50 particles will receive stellar ejecta. The stellar feedback is tightly coupled to our adopted chemical evolution model, which includes both SNe Ia and SNe II with yields from Kobayashi et al. (2000) and Tsujimoto et al. (1995), respectively. Exploding SNe are computed stochastically using a continuous IMF sampling scheme (Revaz et al. 2016). Thus here, a thermal energy equivalent to 4.5×10^{51} erg/SN is released into the ISM, following a blast-wave-like feedback scheme (Stinson et al. 2006) with a 5 Myr delayed cooling time. While GEAR does not include artificial metal diffusion, we use the smooth metallicity scheme to mix the metal-enriched gas effectively (as in GADGET-3; see Section 3.2.5).

3.2.7. GIZMO

GIZMO is a mesh-free hydrodynamics code (Hopkins 2015), a descendant of GADGET-3 in which a kernel-based partition scheme is used to discretize the domain in a set of unstructured "cells" that are allowed to move and reshape with time. The Riemann problem is solved across the effective faces shared by neighboring cells, similarly to what is done in grid-based codes. The version used for this work includes the common cooling and star formation models described in Paper II while stellar feedback is based on the mechanical feedback model described in Hopkins et al. (2018), i.e., both kinetic and thermal energy are distributed among gas cells lying within each star particle kernel according to the evolutionary stage of the SN blast wave (energy or momentum conserving). The SN rate used in this work is described by a piecewise function, where we assume the decaying power-law fit in Lupi et al. (2020) for star particles older than 5.089 Myr, and a constant rate equal to the power-law maximum value for younger stars, aimed at modeling early feedback by massive stars. For consistency, the integrated number of SN events is normalized to ensure one SN per 91 $M_{\odot}$, while the injected energy is set to 5×10^{51} erg/SN in order to reproduce the desired stellar mass at $z = 4$.

4. Common Runtime Parameters

We describe here our choices of common runtime parameters, such as numerical resolution. They are based on what we used in the dark-matter-only cosmological test for a galaxy-sized halo (Section 5 of Paper I), and in the isolated disk test for a Milky Way–sized halo (Section 4 of Paper II).

For the particle-based codes CHANGA, GADGET-3, and GEAR, a spline kernel is used to soften gravity (e.g., Equation (A1) of Hernquist & Katz 1989). The gravitational softening length $\epsilon_{\rm grav}$ in the highest-resolution region is set to 800 comoving pc until $z = 9$, and 80 proper pc afterward. While this resolution is better than what Equation (15) of Power et al. (2003) proposes (∼220 pc), it is used to match the resolution of Paper II at which our fiducial subgrid physics models were initially calibrated. For particles in the lower-resolution region at a corresponding MUSIC level ℓ, the softening length is set at $80 \times 8^{(\ell_{\rm max}-\ell)/2}$ proper pc after $z = 9$, as Power et al. (2003) suggest $\epsilon_{{\rm grav},\ell} \propto N_{200}^{-1/2} \propto (m_{{\rm DM},\ell})^{1/2}$. For particle-based codes, we also require that the minimum hydrodynamical smoothing lengths for gas particles be $0.2\,\epsilon_{\rm grav}$. The exact choice of smoothing scheme is left to each code group's discretion (see Section 5 and Appendix C of Paper II).

Meanwhile, the finest cell size of the mesh-based codes ART-I, ENZO, and RAMSES is set to 163 comoving pc, or 12 additional refinement levels for a 128^3 root resolution in a $(60\ \text{comoving}\ h^{-1}\ \text{Mpc})^3$ box. A cell is adaptively refined into eight child cells on particle or gas overdensities of 4. Given the differences in refinement algorithm among the codes, the parameters that control the overall mesh structure and the aggressiveness of the refinement are left for each code group to decide (see Section 5 of Paper II). These differences can have an impact on the gas density and temperature distributions without stellar feedback (as shown in Sections 5.1 and 5.2), but the impact becomes marginal once stellar feedback is activated (Section 5.4). Further analyses of such differences in the evolution of primordial gas at high z will be presented in future papers from the AGORA Collaboration.

Lastly, we recommend that each group store simulation outputs at 200 epochs.[42] An explicit list of this AGORA-recommended output interval is publicly available, and can be used by anyone to compare their simulation with AGORA.

5. Physics Calibration Steps

Before proceeding to generate the final cosmological simulations, all participating code groups have been asked to complete four rigorous calibration steps. The main objective of these calibrations is to reduce the number of free parameters and artifacts in each code that can have an impact on the evolution of simulated galaxies, which are not valid physical assumptions about structure formation. By adding one physical process at a time into our cosmological zoom-in simulation, we seek a situation where all code groups converge to a final simulation with similar global properties (e.g., similar stellar

[42] Two hundred epochs from $a = 0.062$ ($z \sim 15$) to $a = 0.325$ ($z \sim 2$), equally spaced in $\log(a)$ with $\Delta \log(a) = |\log(384/2013)/200|$, plus a set of redshift snapshots at $z = 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2$. Downloadable at http://physics.snu.ac.kr/cosmo/agora/output_z_cosmorun.txt.

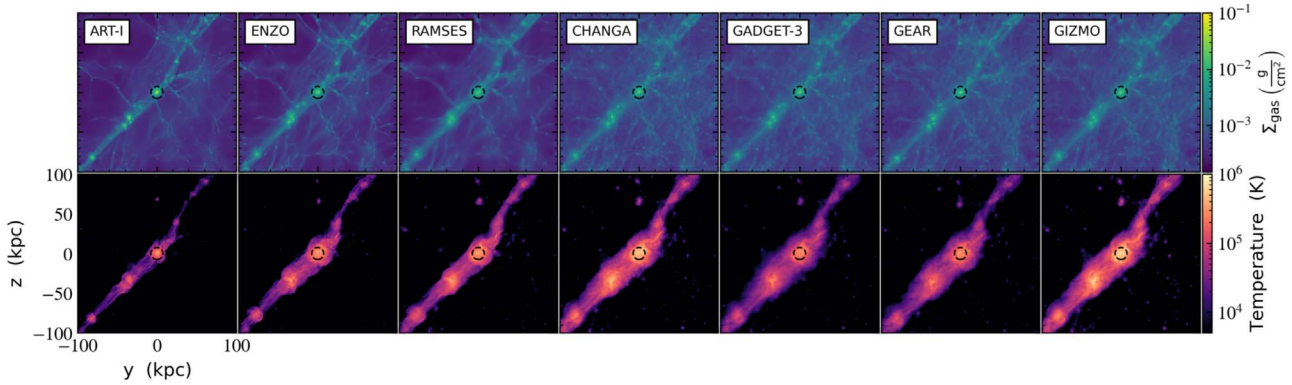

Figure 2. Gas density projection (top) and density-weighted temperature projection (bottom; each made through a slab of thickness 200 kpc) at $z = 7$ from the first calibration step, `Cal-1` (adiabatic evolution test). We indicate the mean R_{200} among the codes (∼7.5 kpc at $z = 7$) with a black dashed circle. The units are proper kiloparsecs. See Section 5.1 for more information on `Cal-1` and this figure, and Section 3.2 for descriptions of participating codes in this comparison. A full-color version of this figure is available in the electronic edition. High-resolution versions of this figure and article are available at the project website, http://www.AGORAsimulations.org/. The simulations are performed by Santi Roca-Fàbrega (ART-I, RAMSES), Ji-hoon Kim (ENZO), Johnny Powell and Héctor Velázquez (CHANGA), Kentaro Nagamine and Ikkoh Shimizu (GADGET-3), Loic Hausammann and Yves Revaz (GEAR), and Alessandro Lupi and Bili Dong (GIZMO).

mass)—and thus, any differences can only be attributed to the chosen stellar feedback prescriptions and intrinsic variations of the codes' numerics. We summarize the calibration procedure with a flowchart in Figure 1.

The first two calibration steps (hereafter `Cal-1` and `Cal-2`) are designed to first acquire qualitative convergence on the main gas properties, by calibrating the gas physics, such as cooling and heating, when star formation is not enabled. In the third calibration step (`Cal-3`), with star formation enabled, but the corresponding stellar feedback disabled, we look for agreement in the main gas properties and in the total stellar mass produced at $z = 7$. Finally, in the fourth step (`Cal-4`), we activate stellar feedback and aim to achieve convergence only in the stellar mass at $z = 4$ to the values predicted by semiempirical models. Each code group is asked to set the feedback prescription to as close to the most widely used one in each code community as possible. This last calibration step is a groundwork from which we can study how galactic properties depend on feedback prescriptions.

An important result of our set of calibrations is that the simulation parameters selected in an isolated disk test (Paper II) cannot be naively used in cosmological simulations, like the ones presented here. Gas properties (e.g., metallicity) and the external radiation field rapidly evolve with redshift, which has a strong impact on gas cooling and thus on star formation. Furthermore, continuous acquisition of fresh gas from the intergalactic medium (IGM) and circumgalactic medium (CGM) makes the cosmological run substantially more complex than that of an isolated disk galaxy.

In this section, we carefully describe the four calibration steps one by one. We start each subsection by explaining the setup, and then go through the important findings and conclusions from each step. One could consider each of our calibration steps as a standalone comparison in itself. Nevertheless, when successively executed and combined with other steps, our calibration procedure provides a solid ground on which advanced cosmological simulations could be performed and trusted. For example, new code groups may test their code's compatibility with other contemporary codes, by following the common initial conditions, the common physics package, and the calibration steps proposed herein.

5.1. Calibration Step One (`Cal-1`): Adiabatic Evolution of Gas

The first calibration step we undertake (`Cal-1`) is designed to detect interplatform variations in the temperature and density of the accreted gas at $z = 7$ when no radiative process or subgrid physics is present. Each cosmological run has been performed without any radiative cooling processes or heating sources, or any subgrid models, such as star formation or a pressure floor. Under such conditions, the system exchanges no energy with its surroundings, and is considered adiabatic. The system's entropy, however, is not necessarily constant as it may increase owing to the presence of shocks. If so, any variation between the codes is in principle caused only by the differences in hydrodynamics solvers—namely, how each code solves the conservation laws of fluid dynamics and how shocks, e.g., in the accreting gas, are captured and treated. Despite the small differences described below, an overall convergence has been found among the seven participating simulation codes.

5.1.1. Findings from `Cal-1`

In Figure 2 we show the projected density (top row) and temperature (bottom row) from `Cal-1` at $z = 7$. The virial radius, denoted as R_{200}, is approximately 7.5 kpc at $z = 7$ across all the codes (see Table 2), shown as black dashed circles. In Figure 2 and similar projection images thereafter, particle-based codes are smoothed using a spline kernel in `yt`.[43] However, these codes are not smoothed in other types of figures and analyses in this paper. Meanwhile, in Figure 3, we show the density–temperature probability distribution function (PDF) at the same epoch for all gas within 100 kpc of the center of the main progenitor. Because the virial radius of the target halo at this redshift is ∼7.5 kpc, we are showing a volume that includes gas not only in the galaxy, but also inside filaments, sheets, knots, and voids.

Overall, the large-scale density structures in all seven panels of Figure 2 are remarkably similar to one another, and the multiphase density–temperature structures in Figure 3 are also

[43] We employ `yt-v4.0`, which better handles SPH particles, an improvement from `yt-v3.3` used in Paper II. See how `yt-v4.0`'s handling of SPH particles differs from that of its predecessors at https://matthewturk.github.io/yt4-gallery/.

Table 2
Global Properties of the Target Galaxy Progenitor in the AGORA CosmoRun Simulation Suite

Code	Redshift z	$M_{200}^{(a)}$ [a] ($10^{10}\,M_\odot$)	$M_\star^{(b)}$ ($10^8\,M_\odot$)	$M_{\rm gas}^{(c)}$ ($10^8\,M_\odot$)	$M_{\rm gas,gal}^{(d)}$ ($10^8\,M_\odot$)	$M_{\rm gas,CGM}^{(e)}$ ($10^8\,M_\odot$)	$\log\,(M_{\star,\rm gal}/M_{200})^{(f)}$
ART-I	8	0.92	0.48	11.36	0.19	11.18	−3.7
	7	1.49	1.04	14.87	0.38	14.50	−3.22
	6	1.83	1.52	17.80	0.56	17.24	−2.86
	5	2.77	1.98	28.50	1.29	27.21	−2.71
	4	13.23	9.22	145.41	21.68	123.72	−2.64
ENZO	8	1.16	0.23	11.03	0.17	10.86	−3.72
	7	1.84	0.43	22.37	0.83	21.54	−3.41
	6	2.26	0.96	30.05	1.58	28.46	−2.97
	5	3.84	2.04	51.41	3.67	47.74	−2.72
	4	16.04	12.72	242.62	58.39	184.23	−2.28
RAMSES	8	1.37	1.21	17.73	2.97	14.75	−2.32
	7	1.84	1.67	19.85	1.51	18.35	−2.51
	6	2.19	2.87	26.59	5.12	21.48	−2.11
	5	3.50	5.12	36.51	10.43	26.08	−1.96
	4	14.79	18.98	139.47	44.32	95.15	−1.97
CHANGA	8	1.43	1.17	29.03	5.94	23.37	−2.26
	7	2.26	2.82	43.22	7.55	35.67	−2.02
	6	2.72	5.09	58.88	17.91	40.97	−1.84
	5	4.15	10.89	72.74	11.76	60.98	−1.68
	4	15.81	39.94	203.04	85.70	117.34	−1.63
GADGET-3	8	1.32	0.48	25.16	5.62	19.54	−2.60
	7	2.17	1.47	38.84	7.41	31.43	−2.26
	6	2.61	4.23	49.25	18.06	31.20	−1.82
	5	4.05	12.75	71.65	26.46	45.20	−1.52
	4	16.15	53.17	216.98	76.24	140.74	−1.51
GEAR	8	1.72	0.67	39.52	8.28	31.24	−2.60
	7	2.52	1.55	58.84	15.51	43.33	−2.33
	6	3.23	3.71	82.14	14.93	67.21	−2.15
	5	4.60	7.77	111.38	40.51	70.87	−1.94
	4	16.34	25.92	286.33	145.52	140.81	−1.88
GIZMO	8	1.12	0.14	10.96	0.0	10.96	−4.24
	7	1.90	0.20	24.56	1.15	23.41	−4.14
	6	2.35	0.92	33.02	0.98	32.04	−3.03
	5	3.65	1.64	41.18	1.32	39.86	−2.86
	4	15.39	36.23	165.59	41.21	124.38	−1.66

Note.
[a] Each column lists the following quantities at the corresponding redshift: $^{(a)}$ total halo mass; $^{(b)}$ stellar mass; $^{(c)}$ gas mass inside the mean R_{200} among codes, where the R_{200} values found are 5.8, 7.5, 8.4, 11.4, and 25.4 proper kpc at $z = 8$, 7, 6, 5, and 4, respectively; $^{(d)}$ gas mass inside the main galaxy or the ISM (which we define as regions with $R < 0.15\,R_{200}$); $^{(e)}$ gas mass in the CGM (which we define as regions with $0.15\,R_{200} < R < R_{200}$); and $^{(f)}$ the ratio of stellar mass (in the main galaxy) to halo mass.

comparable. Unsurprisingly, in both plots, the convergence is very good qualitatively for the particle-based codes CHANGA, GADGET-3, GEAR, and GIZMO as they share gravity solvers and take similar SPH approaches. The three mesh-based codes, ART-I, ENZO, and RAMSES, show minor differences but overall agreement, too. Larger discrepancies are observed when comparing particle-based codes with mesh-based codes. In particular, the differences in the resolved structures in low-density regions at high redshift were discussed in the previous AGORA comparison with dark-matter-only simulations. It is because the particle-based codes achieve better resolution at early times than the mesh-based codes assuming little or no adaptive refinement for the mesh-based codes at high z (for a detailed discussion, see Section 5.3.2 of Paper I). We also notice in Figure 3 that the highest densities that each code reaches are somewhat different, particularly among the mesh-based codes. This is due to the differences in refinement strategies adopted in each code, and we plan to study this issue further in future publications.

5.1.2. Comments on Differences in the Warm–Hot IGM in Cal-1

From Figure 2, one however notices some discrepancies in the temperature maps. While all codes reproduce the virialized hot gas expected around massive halos, with temperatures between 10^5 and 10^6 K, it is clear from Figure 3 that the extension of this hot component to lower densities—the warm gas that surrounds the main galactic systems—slightly differs. In particular, in ART-I, the intergalactic warm gas extends only up to the virial radius indicated in Figure 2 by the black dashed circles, while it extends beyond the virial radius and

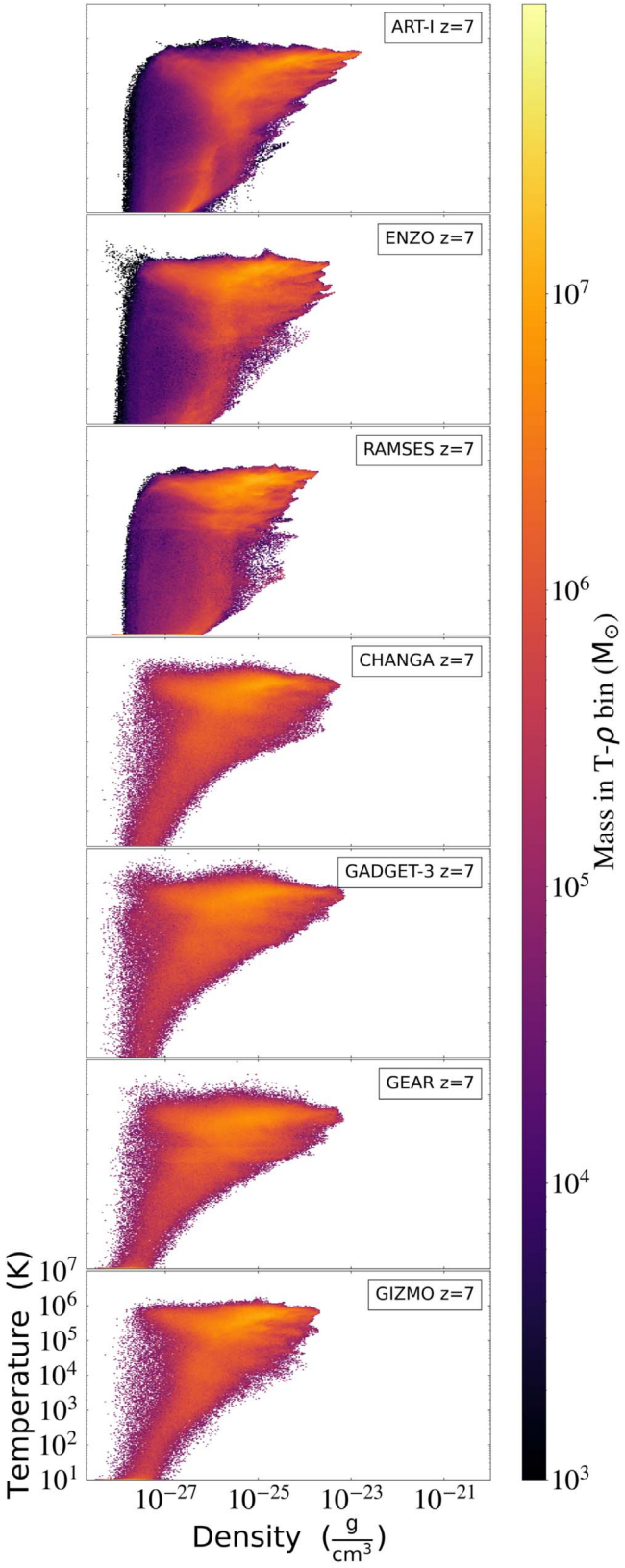

Figure 3. The $z = 7$ composite of the two-dimensional PDF of density and temperature for the gas within 100 kpc of the center of the main galactic system in the `Cal-1` runs. The 100 kpc radius sphere encloses the main galaxy, the CGM, and the nearby IGM. Colors represent the total gas mass in each two-dimensional bin. In all analyses for particle-based codes hereafter—except graphical visualizations, such as Figures 2 and 10—the raw particle fields are used, not the smoothed fields built by `yt`. See Section 5.1 for more information on `Cal-1` and this figure.

encompasses more mass in CHANGA, GADGET-3, GEAR, and GIZMO.

The effects that accretion shocks have on the warm gas around the main galactic systems could be different between codes, as they can be caused by small differences in numerical techniques. This phenomenon has been documented by many authors: (1) Gas could be overheated via collisional heating with dark matter particles due to differences in gravity solvers, integrators, timestepping strategies for force calculations, and refinement strategies (Springel 2010; Lukić et al. 2014; Jia et al. 2020). (2) Gas could be overheated also by the artificial viscosity in sharp accretion shocks in particle-based codes (Scannapieco et al. 2012; Taylor & Miller 2012; Hosono et al. 2016). (3) Gas could be overcooled in accretion shocks due to low resolution in the insufficiently refined CGM (Hubber et al. 2013). We present the first analysis here, but this will be better characterized in a future paper from the Collaboration.

5.2. *Calibration Step Two (`Cal-2`): Cooling and Heating of Gas by a Common Physics Package*

The second calibration step (`Cal-2`) is designed to check if the common physics package (i.e., cooling, heating, and UVB) by GRACKLE V3.1.1 is properly interfaced in all the codes for cosmological runs. Here, each run is performed with GRACKLE V3.1.1 but without any subgrid models, such as a pressure floor, star formation, or feedback. This approach allows us to check the agreement on the gas distribution in the density–temperature plane (expected when the radiative gas physics is treated via the common package GRACKLE V3.1.1), and if all codes use the same initial metallicity.

5.2.1. *Findings from `Cal-2`*

`Cal-2` has turned out to be a critical calibration step, during which the participant code groups have found and fixed problems in their GRACKLE V3.1.1 interface.[44] Note that an earlier version of GRACKLE was implemented and tested for an isolated galaxy disk simulation for all codes (see Section 3.1 of Paper II), but not for a fully cosmological zoom-in run with an expanding simulation volume.

The gas mass distribution from `Cal-2` in the density–temperature plane is shown in Figure 4 at $z = 7$. Since the virial radius of the target progenitor at $z = 7$ is $\sim$7.5 kpc (see Table 2), and we include all the gas inside a sphere of 100 kpc centered on the main halo, the plot includes not just the galactic gas, but most of the IGM inside the Lagrangian zoom-in region. Above $\sim 10^4$ K, the gas cools extremely efficiently owing to both hydrogen and helium recombination. Below $\sim 10^4$ K, however, the cooling of the low-metallicity primordial gas (see Section 2) is very weak due to the absence of efficient cooling channels other than primordial molecules. On the other hand, the low-density gas is strongly heated by the UVB up to $\sim 10^4$ K, while at higher density above the UV self-shielding limit, the gas is heated by adiabatic compression. The combination of these effects leads to the bulk of the gas being found in a well-defined plateau at $\sim 10^4$ K, extending up to high densities (10^{-20} g cm^{-3}).

[44] During our comparison study using an earlier version of GRACKLE, we found that a small correction on the cooling and heating rates was needed in the GRACKLE/CLOUDY tables, to ensure correct gas evolution at high redshift. This issue has been addressed in the GRACKLE V3.1.1 release.

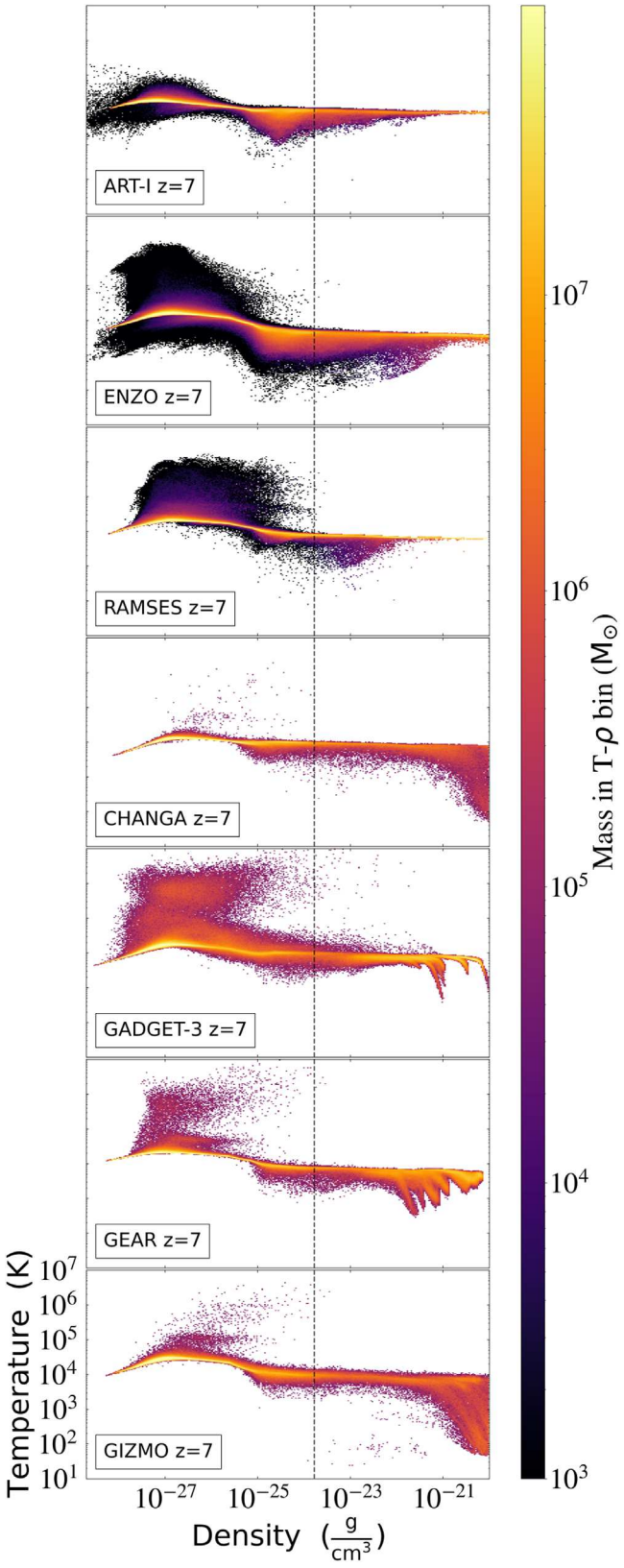

Figure 4. The $z = 7$ composite of the two-dimensional PDF of density and temperature for the gas within 100 kpc of the center of the main galactic system in the `Cal-2` runs (cooling and heating test). The 100 kpc radius sphere encloses the main galaxy, the CGM, and the nearby IGM. Colors represent the total gas mass in each two-dimensional bin. A vertical black dashed line is placed at the value of the star formation density threshold (Section 3.1) to be later adopted in the final simulations in Section 6. See Section 5.2 for more information on `Cal-2` and this figure.

Despite good general agreement in the reproduction of this plateau, discrepancies between the participant codes have been noted. These discrepancies reside primarily in the low-density, high-temperature gas in Figure 4. First, it is worth noting that the mesh-based codes sample the low-density gas with a large number of bins with small mass per bin (blue bins)—which is hard to reproduce by particle-based codes with a (roughly) constant particle mass. In Figure 4, this leads to a large blue area at density 10^{-27}–10^{-24} g cm^{-3}, above and below the 10^4 K plateau. This area is absent in the particle-based simulations. Second, a discrepancy exists in the prediction of the rarefied and shocked gas surrounding the halo and filaments. While the particle-based codes predict the presence of virialized hot gas at 10^{5-6} K (low-density, high-temperature gas at [$\sim 10^{-27}$ g cm^{-3}, $\sim 10^{5-6}$ K] in Figure 4, or a similar gas structure in Figure 6 or 10), it is almost absent in the mesh-based codes. We have carefully studied the behavior of this warm–hot gas, and found that the hot gas is outflowing, while the warm gas is inflowing, confirming that the warm gas surrounding the main galactic system contains shock-heated gas. While at this stage of our analysis, the exact origin of the temperature discrepancies between the codes remains unclear, we hypothesize that they result from the different hydrodynamic schemes adopted (differences in the hot virialized gas have already been mentioned in `Cal-1`), and in particular from how the schemes treat shocks in strongly cooling gas phases.

Finally, it is worth noting that although those discrepancies may look important, they typically disappear as soon as stellar feedback is activated (Section 5.4). Since they have little impact on star formation in our final `CosmoRun` simulations (Section 6), we have chosen to defer detailed discussion to a future paper. Extensive studies on the differences in numerical approaches, and how they manifest themselves in the discrepancies in the warm gas surrounding the main galactic system will be presented in a forthcoming paper by the collaboration (AGORA Collaboration et al., in preparation).

5.2.2. *Comments on the Cooling Tails at High Density in* `Cal-2`

In Figure 4, a repeating pattern of cooling "tails" appears at high density ($\gtrsim 10^{-22}$ g cm^{-3}), especially in the particle-based codes GADGET-3 and GEAR—although we have confirmed that these features also exist in CHANGA, GIZMO, and the mesh-based codes (e.g., RAMSES) at other epochs. After carefully checking the physics in each of the participant codes, we have found that such features are caused by the cooling and heating tables in our common physics package GRACKLE V3.1.1. To illustrate our finding, in Figure 5 we show the tabulated rates of primordial cooling and heating at $z \sim 7$ from our adopted CLOUDY table (see Section 3.1). Here, it is easy to notice how the precomputed table is binned in density and temperature. Readers may notice a larger bin size in the density axis, and that the discrete jumps at high density ($\gtrsim 10^{-22}$ g cm^{-3}) in the cooling and heating rates exactly coincide with the cooling tails in Figure 4. We therefore conclude that the observed cooling tails originate from the density binning in the precomputed CLOUDY table and the differences among the participating codes are due to variations in how exactly each code's cooling and heating solver interfaces with GRACKLE V3.1.1 and its interpolation scheme.

While the cooling tails are an interesting observation, we note that these artificial features have little impact on the final cosmological runs presented in Section 6, because they occur at

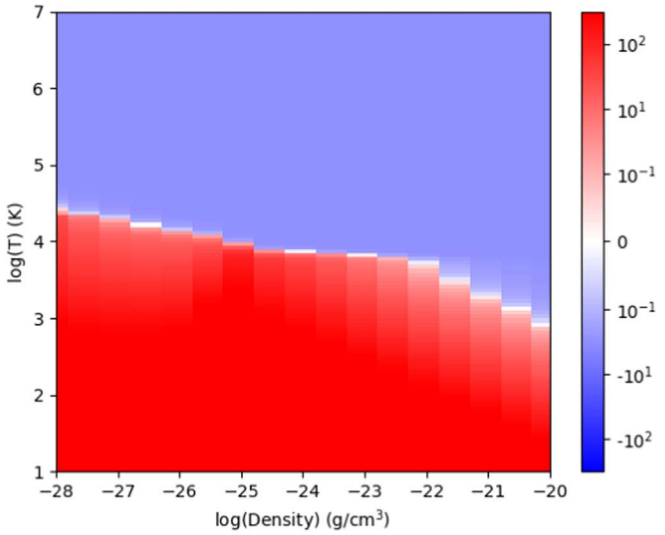

Figure 5. The density and temperature plane colored by the ratio of (heating rate – cooling rate) to cooling rate in each bin, obtained from the CLOUDY table at $z \sim 7$ in GRACKLE V3.1.1.

densities much higher than the star formation threshold, $n_{\rm H,thres} = 1\ {\rm cm}^{-3}$, where, in addition, the pressure is dominated by the artificial pressure floor (Section 3.1). The features start to disappear once the dense gas is consumed by stars at later times (Section 5.3), and will completely vanish as soon as stellar feedback and the pressure floor are activated (Section 5.4).[45]

5.3. Calibration Step Three (`Cal-3`): Common Star Formation Physics

The third calibration step (`Cal-3`) is designed to detect and study the impact of any discrepancies in the implementation of the common star formation prescription (see Section 3.2). Each simulation has been carried out with GRACKLE V3.1.1 and with common star formation and pressure floor prescriptions, but without any stellar feedback. The main objective of `Cal-3` is to ensure that our final cosmological simulation entries in Section 6 are not dominated by variations (or errors) in how the common star formation physics is implemented in each code. At the end of `Cal-3`, each code group confirms that the feedback-free simulations converge within 0.5 dex in stellar masses at $z = 7$, and in stellar mass growth history down to that point.

5.3.1. Findings from `Cal-3`

In Figure 6, we plot the two-dimensional density–temperature PDF at $z = 7$. It displays good agreement on the general features in the density–temperature plane, such as on the shape of the $\sim 10^4$ K cooling plateau, where most of the gas mass resides. However, there are differences, some of which have been discussed in previous sections—e.g., a large number of bins with small mass in the low-density, high-temperature region (blue bins; Section 5.2.1), and the cooling tails at high density (Section 5.2.2). An interesting new discrepancy in Figure 6 is the presence of high-density, low-temperature gas in ART-I, CHANGA, and GIZMO, with its density near the star formation threshold and its temperature near the CMB floor. This artificial feature results from using a stochastic star formation recipe and a particular pressure floor implementation (see Sections 3.2.1 and 3.2.4);[46] however, the discrepancy becomes largely marginal once stellar feedback is turned on as we will discuss in Section 5.4 and Figure 14.

Nevertheless, on the whole, the `Cal-3` entries from the participating code groups exhibit robust overall convergence in the gas distribution around the target progenitor galaxy, as illustrated in Figures 7–9. In Figure 7, we display the gas mass distribution as a function of gas density, including all the gas inside the virial radius R_{200} ($\sim$7.5 kpc at $z = 7$). We find that all participant codes produce a very similar gas density probability distribution inside R_{200}. Note that the convergence is better than that in our disk comparison (Figure 18 of Paper II), in which, by design, gaseous halos—low-density tails toward the left side of this plot—existed only in mesh-based codes, not in particle-based codes. In Figure 8, we show the spherically averaged gas density as a function of radius, again demonstrating solid convergence aside from small variations due to the halo substructures and clumps.[47] In both Figure 7 and Figure 8 we include the fractional deviation from the mean of these profiles to better illustrate the convergence among the codes.

The most relevant result from `Cal-3` is, however, the convergence in the stellar mass $M_\star$ evolution (in a 100 kpc sphere centered at the target progenitor) in Figure 9. Though small variations exist, all codes follow similar stellar mass growth histories, within half a dex of one another at all times. Differences among the codes are due to variations in how the common star formation prescription is implemented (e.g., stochastically in ART-I, CHANGA, GADGET-3, and GEAR versus deterministically in ENZO and RAMSES; see Section 3.1), in the refinement strategy (Section 4), and/or in the numerical accuracies of the hydrodynamics solvers (Section 5 of Paper II).[48]

5.3.2. Comments on Differences in Galactic Morphology in `Cal-3`

Finally, a detailed comparison of the gas and stellar distribution in real space is shown in Figures 10 and 11. In Figure 10 we show the projected density (top row) and temperature (bottom row) of all the gas inside the $(200\ {\rm kpc})^3$ volume (compare with Figure 2 in `Cal-1`). The mean virial radius R_{200} among the codes is shown as a black dashed circle. In the gas density map, the large-scale structures are nearly identical across all participant codes, although the aforementioned differences in the low-density region (discussed in

[45] Although this feature does not affect the final simulations presented herein, we suggest GRACKLE users exercise caution when using the default CLOUDY tables provided with the package. A new table with smaller density bins and/or a careful interpolation scheme would be needed, if they are interested in studying very dense gas when no star formation is present.

[46] The ART-I code, for example, uses stochastic star formation along with a treatment to avoid complete gas depletion in a star-forming gas cell (see Section 3.2.1). Hence, after a cell spawns a star particle, a fraction of gas is still left in the cell with the same temperature as before but with a significantly lowered density. Due to the imposed pressure floor, equilibrium with the surrounding cells can only be achieved through rapid cooling and a slight increase in density. This process results in a buildup of the observed cold gas near the CMB floor. Similar features have been reproduced in other codes (e.g., RAMSES) when stochastic star formation is employed.

[47] The profile center is set to be the location of maximum stellar density within a successively shrinking distance from the dark matter center of mass.

[48] Note that ENZO produces 2–3 times fewer stars than the other codes. Unlike the other codes, the only tunable parameters in ENZO's star formation module are the star formation efficiency and the density threshold, both fixed in this work (see Section 3.1). Thus, it has been difficult to further adjust ENZO's star formation to acquire better convergence.

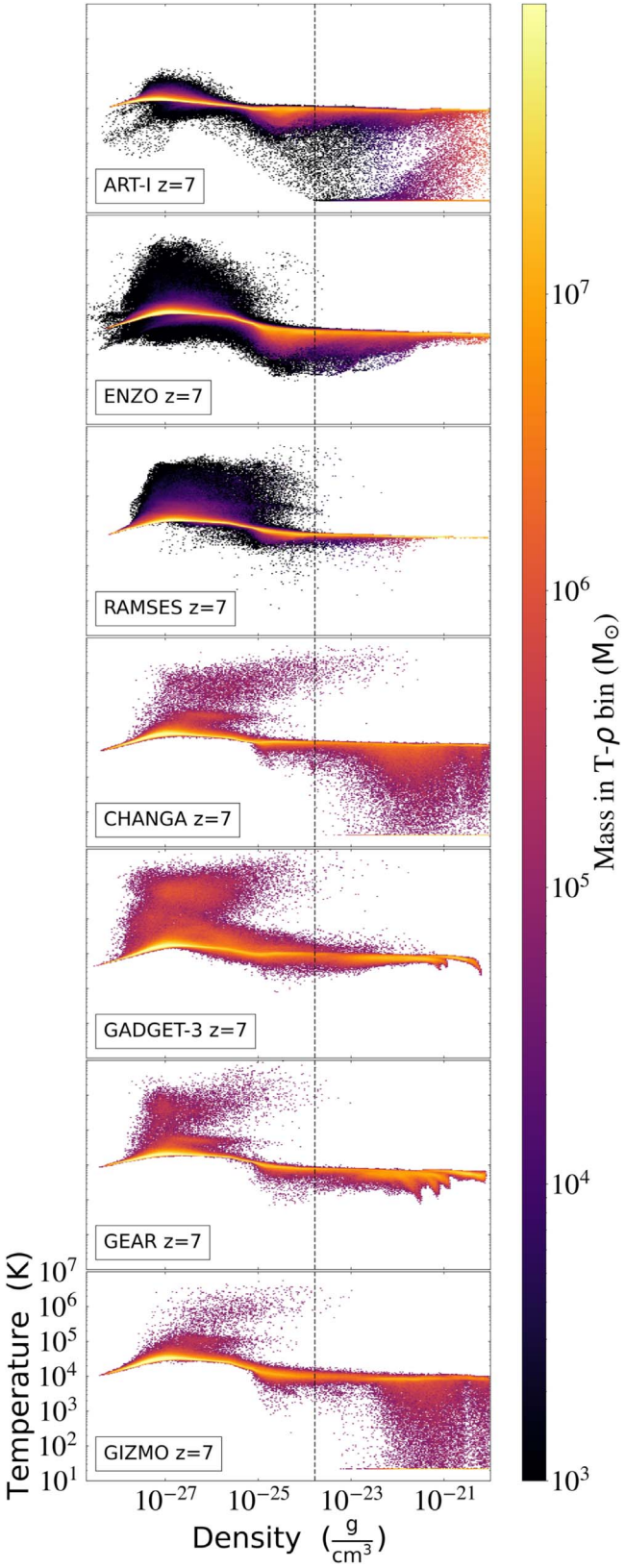

Figure 6. The $z = 7$ composite of the two-dimensional PDF of density and temperature for the gas within 100 kpc of the center of the main galactic system in the `Cal-3` runs (star formation test). The 100 kpc radius sphere encloses the main galaxy, the CGM, and the nearby IGM. Colors represent the total gas mass in each two-dimensional bin. A vertical black dashed line marks the density threshold for star formation. See Section 5.3 for more information on `Cal-3` and this figure.

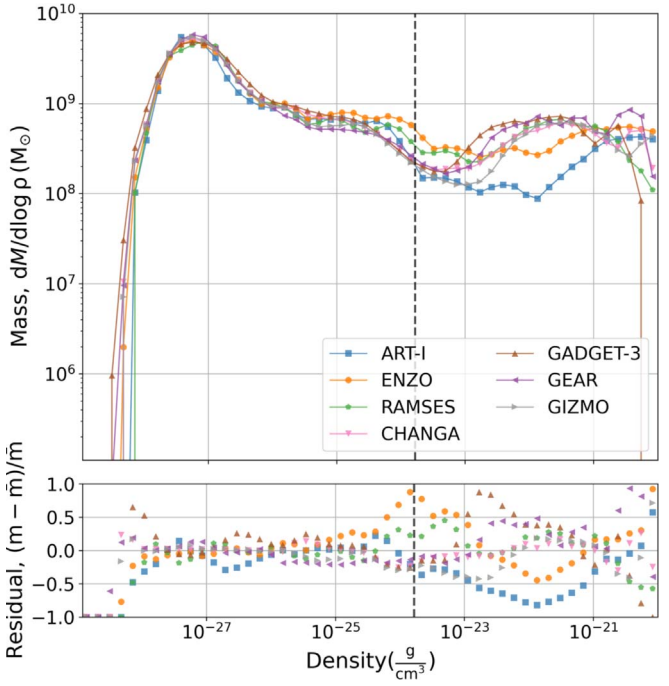

Figure 7. Distribution of gas mass as a function of gas density at $z = 7$ for all the gas inside the target progenitor's mean R_{200} (∼7.5 kpc at $z = 7$) in `Cal-3`. The vertical dashed line denotes the star formation threshold, $n_{\rm H,thres} = 1$ cm^{-3}. Shown in the bottom panel is the fractional deviation from the mean of these profiles. See Section 5.3 for more information on `Cal-3` and this figure.

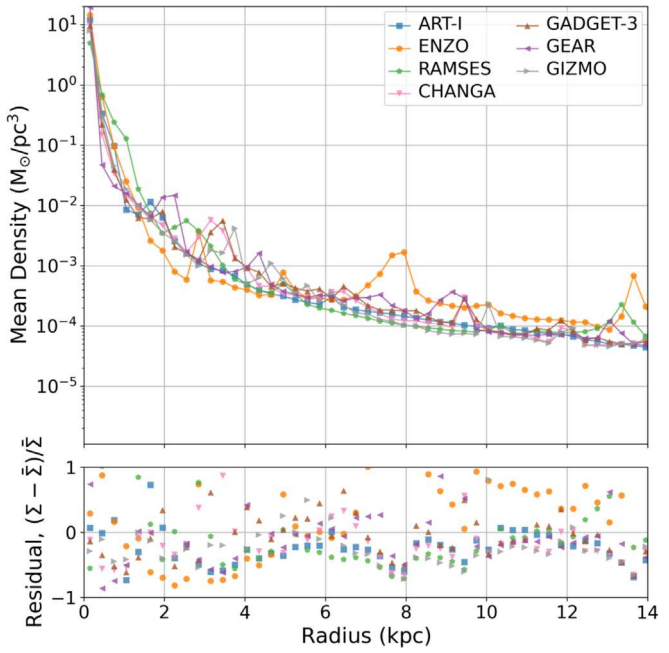

Figure 8. Spherically averaged gas density profiles as functions of distance from the galactic center at $z = 7$ for the `Cal-3` runs. Shown in the bottom panel is the fractional deviation from the mean of these profiles. See Section 5.3 for more information about how the center of the system is selected, the `Cal-3` runs, and this figure.

Sections 5.1.1 and 5.2.1 with `Cal-1` and `Cal-2`, respectively) still exist between the mesh-based and particle-based approaches. Figure 11 demonstrates this more dramatically, in which we show the projected gas density (top row),

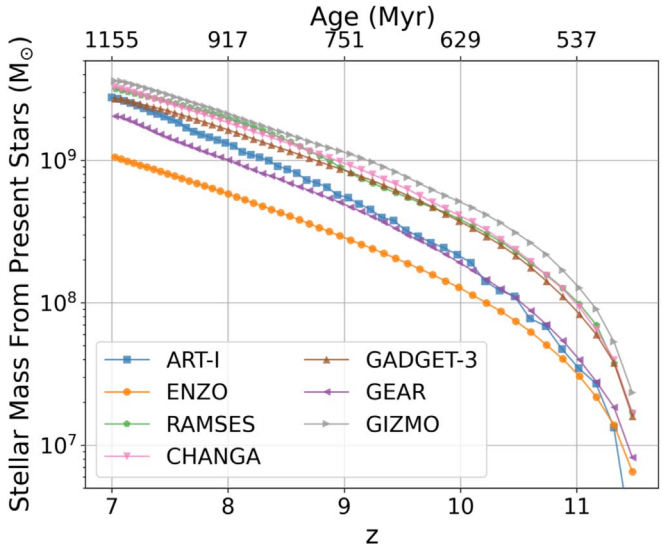

Figure 9. Stellar mass growth histories for the `Cal-3` runs in a 100 kpc sphere centered at the target progenitor. The curve is computed using the ages or creation times recorded in the star particles at $z = 7$.

temperature (middle row), and stellar surface density (bottom row) at $z = 7$ inside a $(4R_{200})^3$ volume. It is notable that, in the stellar surface density map, the particle-based codes harbor more satellites (clumps of star particles) than the mesh-based codes. This discrepancy is caused by the same effect that leads particle-based codes to preserve more substructures in the low-density region. It has been well documented that due to the lack of force resolution at high z, mesh-based codes tend to suppress the low-mass end of the halo mass function (see Section 5.1.1 of this article, or Section 5.3.2 in Paper II).

Also in Figure 11, differences exist in the temperature map between the mesh-based and particle-based codes, particularly in the regions next to the galaxies and filaments. This difference manifests itself as a diverging distribution in the density–temperature PDF at $\sim 10^{-27}$ g cm^{-3}, $\sim 10^{5-6}$ K in Figure 6. We recall, however, from Section 5.2.1 that the observed temperature differences become irrelevant as soon as stellar feedback is activated, and thus have little impact on the results of the final zoom-in cosmological runs (`CosmoRun`) in Section 6.

5.4. Calibration Step Four (`Cal-4`): Favorite Stellar Feedback Prescription of Each Code

The objective of this last calibration step (`Cal-4`) is to get convergence on the stellar mass of the main progenitor at $z = 4$ within 0.5 dex, to the value predicted by semiempirical models based on abundance matching techniques (e.g., Rodríguez-Puebla et al. 2017). The main motivation for `Cal-4` is to come up with a realistic simulation resembling observed galaxies, by adopting each code group's "favorite" feedback—the closest to the most widely used one for research in each code community. Each code group's cosmological simulation has been carried out with Grackle v3.1.1, a common star formation prescription, and its own choice of stellar feedback and metal production (see Table 1 and Section 3.2). Each group has been asked to provide a reference with detailed information on its favorite feedback prescription (as in Section 3.2). Although this part is time-consuming, at the end of `Cal-4` we establish a common ground based on which we can compare the effects of each group's favorite feedback on the evolution of galaxies and the CGM.

5.4.1. Calibration Target in `Cal-4`

According to the predictions by the aforementioned semi-empirical models, the expected stellar mass inside the main galactic system of an $M_{200} = 2 \times 10^{11} M_\odot$ halo at $z = 4$ is $\sim$1–1.5 $\times 10^9 M_\odot$. Since our selected halo (see Section 2) experiences a relatively violent assembly history by $z = 4$, we have extended the target range of the stellar mass $M_\star$ to $\sim$1–5 $\times 10^9 M_\odot$ at $z = 4$. The width of the target mass range is to allow flexibility when each code group selects its stellar feedback scheme. `Cal-4` requires the greatest amount of time among all the calibration steps. Typically, the process is not over with a single simulation, but requires several iterations carried out by each participating code group. The simulations the groups acquire after these iterations become the final entries in Section 6 (named `CosmoRun`). In this subsection we briefly discuss only the calibration process in `Cal-4` and do not yet present a detailed analysis of each code group's final simulation entry—the latter will be discussed in full detail in Section 6.

5.4.2. Findings from `Cal-4`

At the end of `Cal-4`, the participating code groups have found a need to use stronger stellar feedback than they commonly use in their communities in order to achieve the target stellar mass at $z = 4$. However, none of them have used unrealistic feedback parameters. In Figure 12 we show the stellar mass growth histories of the final simulation entries. Each curve has been obtained using the star particles residing inside an R_{200} sphere centered on the target progenitor galaxy at $z = 4$. Therefore, Figure 12 shows the stellar mass assembly history (SMAH) inside R_{200}, not the star formation history (SFH) of the main galactic system; thus it shows only an upper limit for the generated stellar mass.[49] The plot demonstrates how all codes successfully converge to the agreed $M_\star$ range, although the SPH codes tend to have higher $M_\star$ at $z = 4$. Comparing Figure 12 with Figure 9, in each code we observe the expected decrease of stellar mass growth due to stellar feedback (notice the change in the y-axis). The shape of the SMAH differs from one code to another because of the different stellar feedback prescriptions implemented in the codes, which can affect star formation differently at a given epoch. The timing discrepancies among the codes in the halo assembly history could also cause differences in the SMAHs. Indeed, the exact timing of a major merger occurring at $z \sim 4$ could precipitate sizable variations in the SMAH, and in the gas and stellar properties discussed in Section 6 (see Sections 6.2 and 6.3 for more discussion).[50] Lastly, readers may notice that the intercode differences are larger at early times (e.g., the variation is $\sim$1.5 dex at $z = 10$ but $\sim$0.5 dex at

[49] Unlike the SFH, the SMAH includes not only the stars formed inside the target progenitor (in situ), but also the stars formed outside and brought in, e.g., by merging satellites (ex situ). In the SMAH, the stellar mass may decrease due to the mass loss that occurs when the galaxy interacts with its neighbors. In future studies, we plan to compare the actual SFH (rather than the SMAH).

[50] The discrepancies in the exact timings of mergers and star formation events could affect the discussion of various galactic properties in Section 6. In particular, at high z, major mergers are common and can violently disturb the gas inside the galaxy and in its CGM by generating shocks and changing the gas distribution in the density–temperature plane. These perturbative events do not occur at the exact same redshift in all codes (see Section 5.3.2 of Paper I), complicating the intercode comparison. In future papers, we will extensively study variations in the participating codes' merger trees.

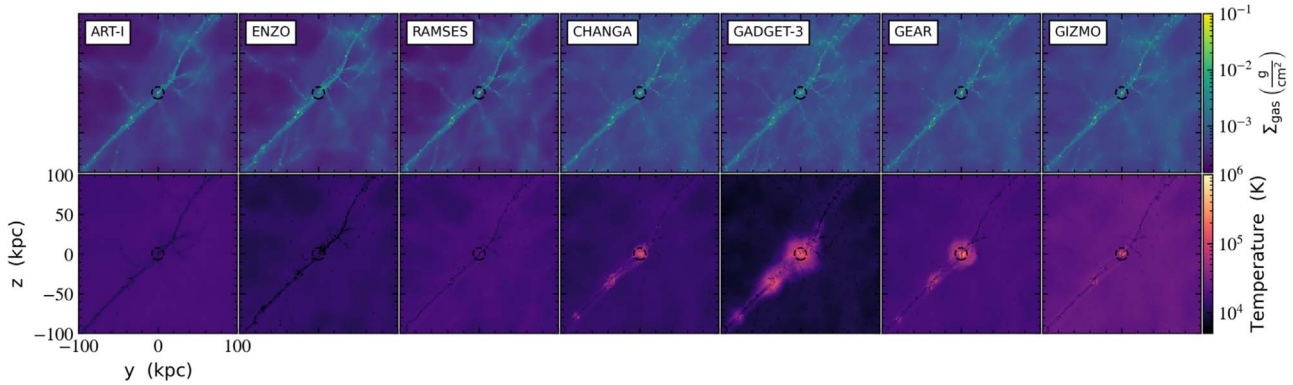

Figure 10. Gas density projection (top) and density-weighted temperature projection (bottom) at $z = 7$ from the third calibration step, `Cal-3` (star formation test). We indicate the mean R_{200} among the codes (∼7.5 kpc) with a black dashed circle. The units are proper kiloparsecs. The projections along the other axes are available as digital supplements to this article. See Section 5.3 for more information on `Cal-3` and this figure.

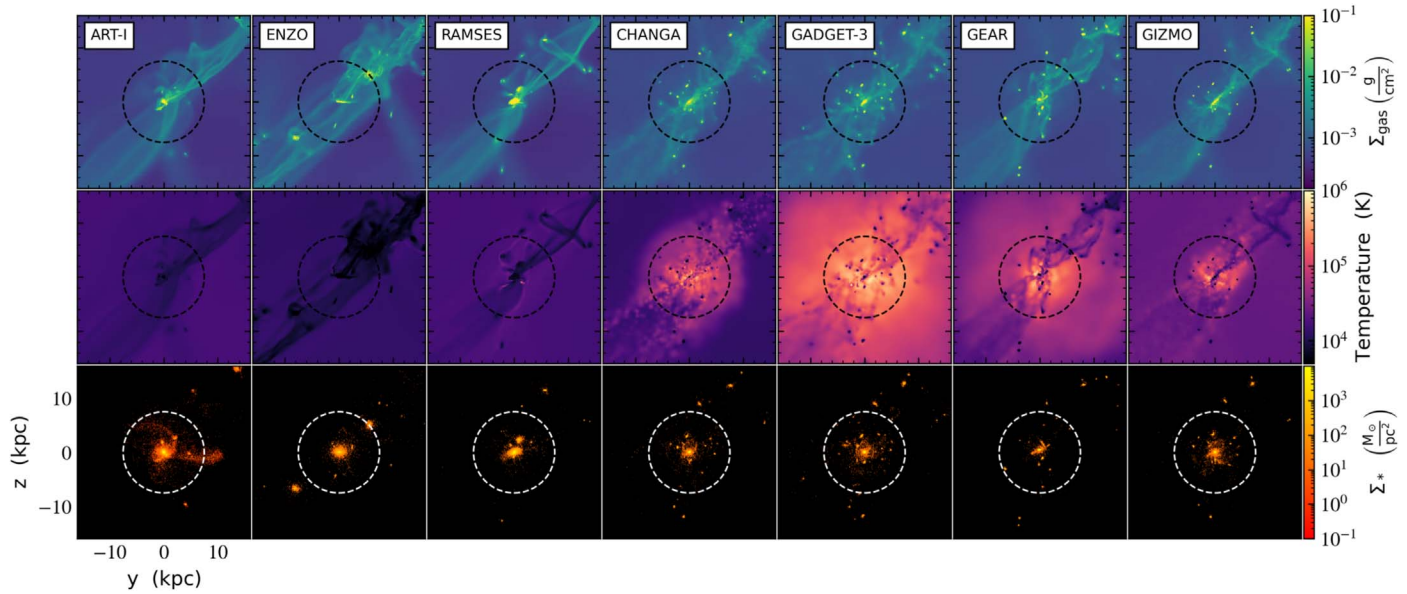

Figure 11. Same as Figure 10, but now in zoomed-in regions. Gas density projection (top), density-weighted temperature projection (middle), and stellar surface density (bottom) at $z = 7$ from the third calibration step, `Cal-3`. The width of each panel is $4R_{200} = 30$ kpc. The mean R_{200} among the codes (∼7.5 kpc) is indicated with a black/white dashed circle. See Section 5.3 for more information on `Cal-3` and this figure.

$z = 4$). Indeed, previous research has found that different stellar feedback implementations can exacerbate the discrepancy at high redshift (e.g., Hayward & Hopkins 2017).

With this final result, we conclude the entire calibration procedure. The code groups that completed the four calibration steps, `Cal-1` to `Cal-4`, have obtained the final `CosmoRun` simulations. In the next section, we present and analyze the properties of these final simulation entries from the code groups down to $z = 4$.

6. The AGORA `CosmoRun` Simulations

In this section, we introduce the AGORA `CosmoRun` simulations acquired from the rigorous calibration steps in Section 5. As we present the analysis of their stellar and gas components, we focus on five redshifts, $z = 8$, 7, 6, 5, and 4.[51] The simulations have been running down to even lower redshift, and the full analysis—the CGM evolution down to, e.g., $z = 2$, in particular—will be presented in forthcoming papers from the AGORA Collaboration.

6.1. Global Properties of the Target Galaxy Progenitor

We start by analyzing the global bulk properties of the target galaxy progenitor in `CosmoRun`. In Table 2 we list the total virial mass, M_{200}, and the gas and stellar masses enclosed inside a sphere whose radius is the mean R_{200} among the codes. We also include the gas masses inside the main galaxy versus those in the CGM (i.e., $M_{\rm gas,gal}$ for $R < 0.15R_{200}$ versus $M_{\rm gas,CGM}$ for $0.15R_{200} < R < R_{200}$), and the stellar-to-halo mass ratio, $M_{\star,\rm gal}/M_{200}$, obtained by using the star particles inside $0.15R_{200}$ (rightmost column in Table 2; see also Figure 13). It should be noted that we do not expect to find perfect convergence in all the properties here, but expect substantial dependence on the stellar feedback prescription adopted by each code group. This dependence will be especially evident in the spatial distribution of gas in and around the target halo, and also in its temperature and metallicity.

[51] 1.09, 1.22, 1.40, 1.63, and 1.96 Gyr in cosmic time, respectively.

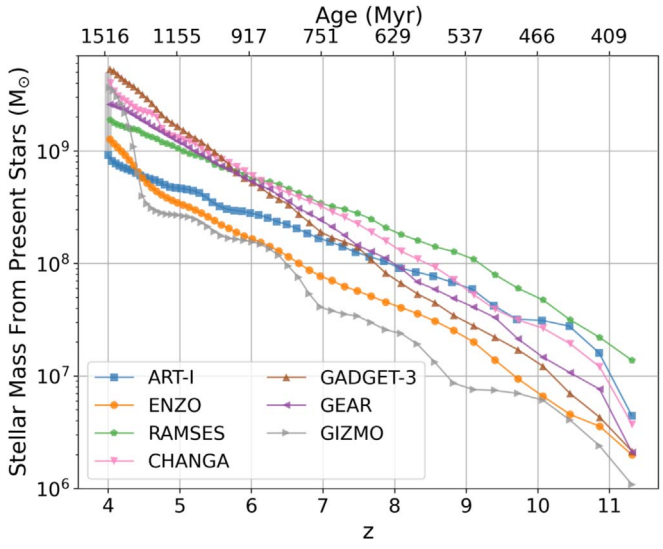

Figure 12. Stellar mass growth histories for the `Cal-4` runs inside an R_{200} sphere centered at the target progenitor. The curve is computed using the ages or creation times recorded in star particles at $z = 4$. The stellar mass range at $z = 4$ targeted in our calibration is $M_\star \sim 1$–$5 \times 10^9\, M_\odot$, as motivated by semiempirical models. What we show here is an upper limit for the total $M_\star$ formed inside R_{200}. It is in Figure 13 where we can make a fair comparison of $M_\star$ formed inside the galaxy with predictions from semiempirical models. See Section 5.4 for more information on `Cal-4` and this figure.

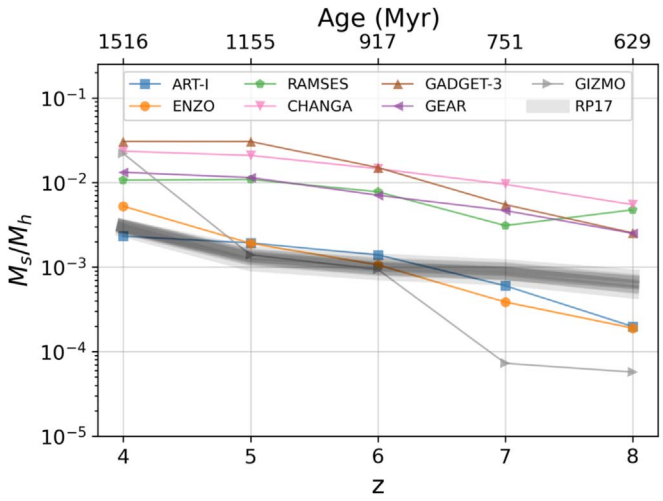

Figure 13. Evolution of the stellar-to-halo mass ratio, $M_{\star,\mathrm{gal}}/M_{200}$, from $z = 8$ to $z = 4$ in the `CosmoRun` simulations (rightmost column in Table 2). Gray shadowed regions indicate the predicted ranges of the ratio by the semiempirical model of Rodríguez-Puebla et al. (2017), obtained using the halo mass at each redshift, in each simulation. See Section 6 for more information on `CosmoRun` (Section 6.1 in particular for more on this figure).

Table 2 illustrates that all the participating codes converge on the stellar and total masses within <0.5 dex of one another. This convergence is not surprising as it is a consequence of the calibration strategy used (`Cal-4`; Section 5.4.1). The small deviations from code to code in the total mass, M_{200}, are due to the timing discrepancies in the halo assembly history (Section 5.4.2 and footnote (50)). On the other hand, the relatively larger deviations in the gas mass inside the virial radius, M_{gas}, or in the ratio of gas masses in the main galaxy to those in the CGM (i.e., $M_{\mathrm{gas,gal}}$ versus $M_{\mathrm{gas,CGM}}$), are a direct consequence of the different stellar feedback strategies adopted. In fact, the strength of the outflows generated by stellar feedback has a strong impact not only on the amount of gas remaining inside the virial radius, but also on how efficiently the cold inflows replenish the galaxy with fresh gas. A detailed analysis of the thermodynamics and kinematics of the gas is made in Sections 6.2 and 6.4.

In Figure 13 we show the stellar-to-halo mass ratios, $M_{\star,\mathrm{gal}}/M_{200}$, in the `CosmoRun`, computed at $z = 8$, 7, 6, 5, and 4 (see also the rightmost column in Table 2), compared with predictions from semiempirical models (e.g., Rodríguez-Puebla et al. 2017). The gray shadowed regions indicate the stellar-to-halo mass ratio obtained from a semiempirical model using M_{200} at each redshift, in each simulation. Since in `Cal-4` we calibrate each simulation's stellar feedback so that the stellar mass produced is in the range of $\sim$1–$5 \times 10^9\, M_\odot$ at $z = 4$ (see Figure 12 and Section 5.4.1), all seven lines do not deviate by more than 1 dex from one another at $z = 4$. In addition, the difference between the simulated stellar-to-halo mass ratios and the semiempirical predictions is less than 1 dex at $z = 4$, because it is designed as such in `Cal-4`. However, the semiempirical predictions lie below the simulated values in most of the codes. The mismatch is due to the fact that our target halo does not have the assembly history of a prototypical halo of $10^{12}\, M_\odot$ at $z = 0$, but that of a halo that assembled early and had a quiescent period from $z = 2$ to 0 (Section 2). This bias yields a higher-than-expected stellar mass at $z \gtrsim 4$. At higher redshift ($z \gtrsim 7$), the differences among the simulated stellar-to-halo mass ratios and those between the simulated ratios and the semiempirical predictions are significantly larger. They are due to the variations in the feedback prescriptions, causing changes in the amount of star-forming gas available at each redshift, and hence in the SFH.

6.2. Gas Properties

Because deviations in stellar feedback are better reflected in gas, gas properties in simulations can be used to compare and calibrate the stellar feedback prescriptions employed. It is not within the scope of this paper to determine which stellar feedback in which code better fits the observations. Instead, we aim to show which gas properties are more sensitive to feedback, and to provide the community with a common ground on which to make new comparisons. In this subsection, we present only a general analysis of the gas properties. This first analysis is currently being extended and the result will be presented in a future paper focused on the evolution of the CGM.

6.2.1. Gas Density and Temperature

The first figure of this subsection, Figure 14, displays the gas density–temperature PDF, which can be compared with Figures 3, 4, and 6 from our calibration steps `Cal-1` to `Cal-3` (see Section 5). Note that, in this plot, we only show the gas inside R_{200} (see the footnote of Table 2), while Figures 3, 4, and 6 include gas out to the IGM. From Figure 14, we see that, once stellar feedback is activated, the convergence we always get is only in the shape of the $\sim 10^4$ K cooling curve. Notable differences between the codes in Figure 14 include the following: (1) The blue bins with small mass per bin in the mesh-based codes reflect very diffuse gas, which is not well represented in the particle-based codes (as discussed in Sections 5.2.1 and 5.3.1). (2) The total gas mass M_{gas} inside R_{200} changes significantly between codes due to the different

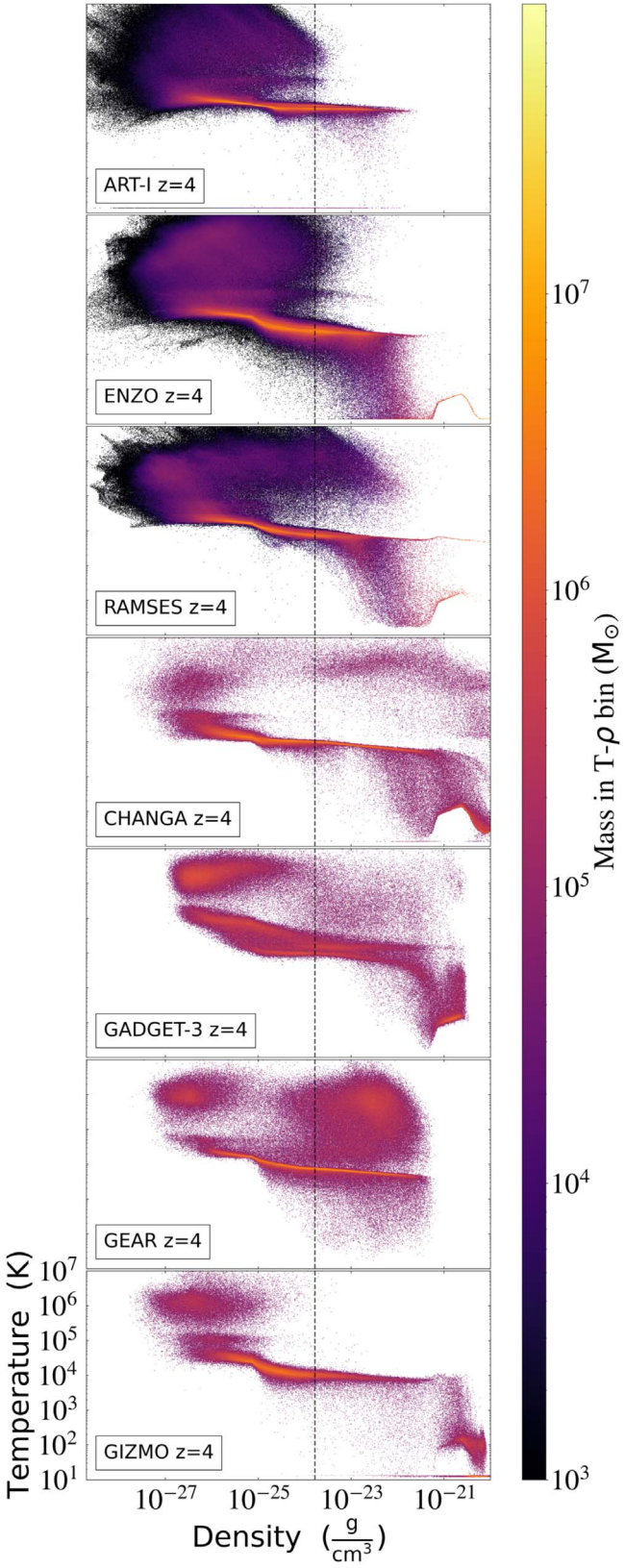

Figure 14. The $z = 4$ composite of the two-dimensional PDF of density and temperature for the gas within the mean R_{200} among the codes (∼25.4 kpc) from the target galaxy's center in the CosmoRun simulations. This is similar to Figures 3, 4, and 6, but unlike those in the previous figures, the sphere of R_{200} here encloses the main galaxy and CGM, but not the IGM. Colors represent the total gas mass in each two-dimensional bin. A vertical black dashed line marks the density threshold for star formation. See Section 6 for more information on CosmoRun (Section 6.2 in particular for more on this figure).

stellar feedback strategies adopted (see Section 3.2) and to the timing discrepancies (see Section 5.4.2 and footnote (50)), for which a clear example appears when comparing the total $M_{\rm gas}$ (number of bins and colors) in R_{200} of, e.g., ART-I and CHANGA (see also Figure 17). The exact timing of a major merger occurring at $z \sim 4$ partly explains the discrepancy in the PDF between different codes. For example, while ART-I still undergoes the merger at $z = 4$, other codes have already experienced it at slightly earlier times (see Section 6.3 and Figure 21). (3) In addition to driving the gas out of R_{200}, the different stellar feedback strategies may instigate other differences in the PDF, particularly in the warm–hot gas phase ($\sim 10^{5-7}$ K) above the threshold for star formation, $n_{\rm H,thres} = 1\ {\rm cm}^{-3}$. Indeed, the gas in star-forming regions is sensitive to variations in the stellar feedback strategies used to release energy and momentum from newly formed stars. In particular, the use of a delayed cooling strategy (in RAMSES, GADGET-3, and GEAR) may result in the accumulation of warm–hot gas in a dense state, around star-forming regions. The superbubble feedback scheme used in CHANGA produces a similar effect on the warm–hot dense gas. (4) Lastly, the cold diffuse gas near the CMB floor, visible only in ART-I, is due to the code's stochastic star formation recipe and its particular pressure floor implementation (as discussed in footnote (46) and Section 5.3.1).[52]

To better illustrate the effect of stellar feedback on the gas in the galaxy, the CGM, and the IGM, we show the evolution of the projected density and temperature in each code in Figures 15 and 16. The mean virial radius, R_{200}, at each redshift (see the Figure 15 caption) is marked with a red/black dashed circle. In these figures, we confirm the differences in the spatial distribution and thermal structure of the gas, due to variations in the stellar feedback strategies, despite the fact that all the participating codes produce similar stellar mass at our target epoch, $z = 4$. Although the differences in gas density and temperature may appear dramatic in Figures 14–16, we find good agreement in the density distribution, especially in the nonextreme-density range. This result can be observed in Figure 17, where we show the evolution of the gas density PDF of all the gas inside R_{200} from $z = 8$ to $z = 4$. We clearly see that most of the codes agree on the total gas mass—the area below the curve—in the intermediate-density range, [$\sim 10^{-27}$, $\sim 10^{-23}$] g cm^{-3}. Obviously, discrepancies in the lowest- and highest-density bins exist, produced by the various reasons in Figure 14 (note that Figure 17 shows the values of Figure 14 integrated along its y-axis).

6.2.2. Gas Metallicity

Metallicity is a good tracer of changes in galactic evolution. The metal content of gas inside the galaxy and its CGM

[52] As a final note to Figure 14, the gas at $\gtrsim 10^{-21}$ g cm^{-3} is seen to be heated up to $\sim 10^2$ K (except in ART-I and GEAR, in which such dense gas is nonexistent for the moment). This heated gas is caused by GRACKLE's redshift-dependent UVB with self-shielding (Section 3.1), and is observed even in a simple one-zone test using GRACKLE. The source of the heating is assumed to be the reemission of absorbed radiation inside the dense gas cloud. The shielded CLOUDY tables are made by integrating into the star-forming cloud for a distance set by the Jeans length at a given density and temperature (with a maximum of 0.1 kpc). Over this length, UVB radiation absorbed by the outer layers of the cloud can be reemitted, causing some heating on the inner layers. We suggest GRACKLE users exercise caution when using the default shielded CLOUDY table provided with the package (e.g., depending on the simulation setup and resolution, one may want to disable the UVB above a certain density).

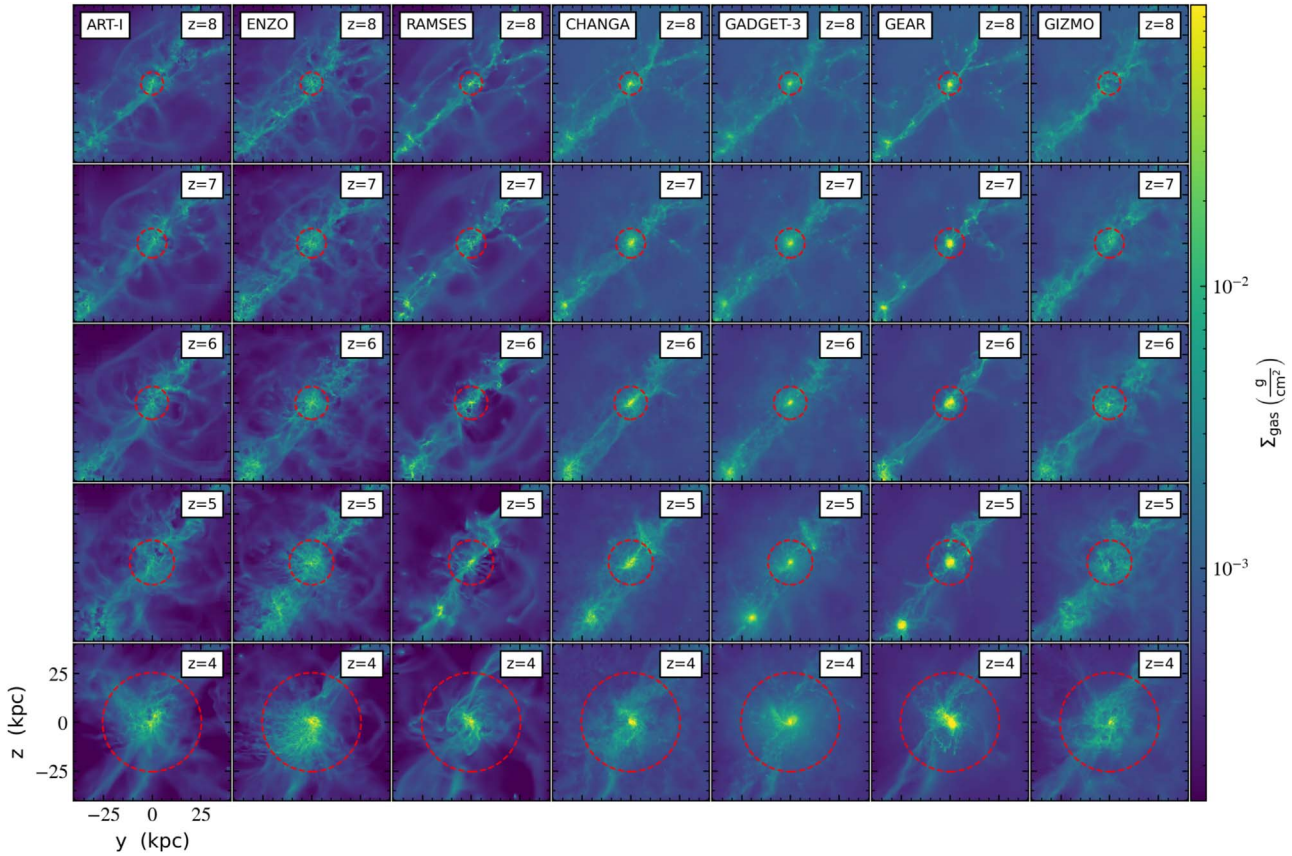

Figure 15. Gas surface densities at $z = 8$–4 from our final CosmoRun simulation suite, centered on the center of mass of stars and dark matter belonging to the target galaxy progenitor. Here and in the following figures we indicate the mean R_{200} among the codes at each redshift with a red dashed circle (5.8, 7.5, 8.4, 11.4, and 25.4 proper kpc at $z = 8$, 7, 6, 5, and 4, respectively). The units are proper kiloparsecs. The projections along the other axes are available as digital supplements to this article. See Section 6.2 for more information on CosmoRun and this figure.

depends on how efficiently outflows remove metal-rich gas from dense star-forming regions. The metal enrichment of the IGM is also dictated by outflows, as the IGM is the recipient of gas pushed out of the virial radius. The exchange of metals between the CGM and IGM also determines the gas evolution in time on the density–temperature plane, as it strongly affects how quickly the gas cools and regulates the interplay between star formation and feedback. Metallicity indeed provides important information on the differences between the feedback schemes employed, and on their ability to fit observations (Suresh et al. 2015; Kacprzak et al. 2019; Lehner et al. 2020).

Before presenting the next figures on metallicities, it is important to remind the reader that all code groups used metal yields in SNe that are similar to the ones in the AGORA common physics (see Section 3.1). Using metal yields similar to the common ones allows us to conjecture that the differences observed in gas metallicity are explained mostly by the variations in stellar feedback—and/or the metal diffusion schemes—presented in Section 3.2. As a consistency check, in each CosmoRun simulation we compute the ratio of the total metal mass to the total stellar mass inside the entire simulation box at $z = 4$ (i.e., the effective metal yields in the fourth column of Table 1). Our calculation confirms that, although each code group is using its favorite metal production strategy, its effective yield value matches what each group assumes in the code's deposit scheme, and is in agreement within less than half a dex of those of the other codes.

First, in Figure 18, we show the projected gas metallicity at $z = 8$, 7, 6, 5, and 4. It is important to mention that a correct interpretation of this figure requires information on the total gas distribution (Figure 15), e.g., most metals in GEAR are in low-metallicity dense gas in the inner parts of the halo. Some codes show high metallicity around the main galaxy (e.g., RAMSES, CHANGA, and GADGET-3), while others exhibit lower values (e.g., ART-I, ENZO, and GIZMO). The former codes are the ones that tend to keep gas and metals around the star-forming regions, while the latter codes are able to push them out to the CGM, or even to the IGM (see also Figure 23). The discrepancy seen here is also due to the spatial distribution of metals being highly sensitive to how efficient the stellar feedback is at driving metal-enriched outflows (see Figure 23), and to how efficient the metal diffusion is at polluting neighboring cells/particles.

We reach a similar conclusion by analyzing the PDF of metallicity and metal mass in Figures 19 and 20, respectively. Here we include only the gas inside a sphere of R_{200} from the target progenitor's center. Figure 19 shows that RAMSES, CHANGA, GADGET-3, and GIZMO exhibit large amounts of high-metallicity ($\gtrsim 1\ Z_\odot$) gas in and around the main galaxy, while ART-I, ENZO, and GEAR show smaller amounts. This difference confirms that the overall gas metallicity distribution depends strongly on the efficiency of stellar feedback. Furthermore, in Figure 20—while the global features in the PDF have been discussed in the section relevant to Figure 14—we find variations in the total metal mass kept inside R_{200}. The

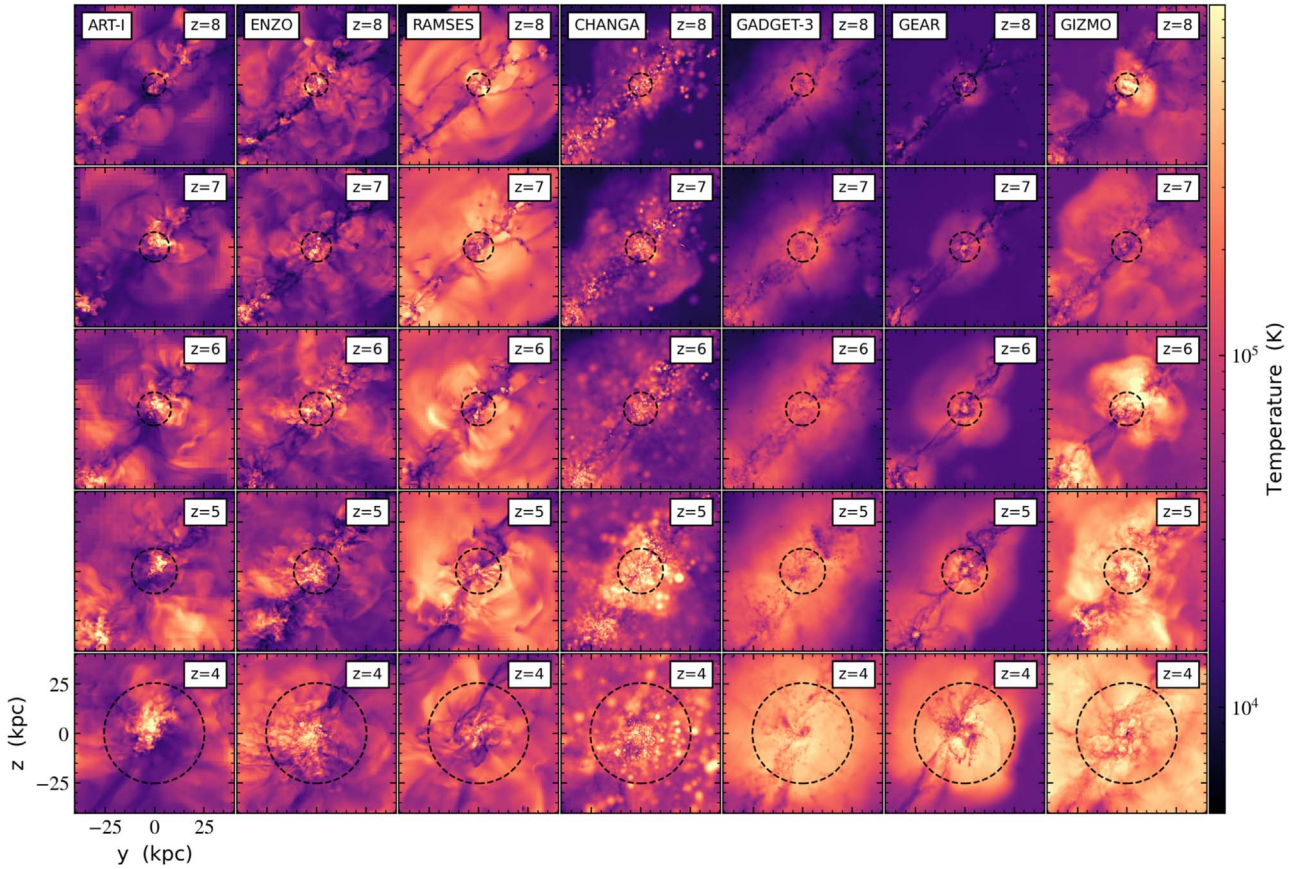

Figure 16. Same as Figure 15, but now showing the density-square-weighted projections of gas temperature in our CosmoRun simulation suite. The units are proper kiloparsecs. See Section 6.2 for more information on CosmoRun and this figure.

stellar feedback in ART-I and ENZO rapidly pushes the metals out to the low-density and low-metallicity gas in the CGM and then to the IGM, leaving only a few dense star-forming regions with high metallicity. In contrast, the other codes keep most of the metals inside R_{200}, showing more regions with high metallicity in the gas density–temperature plane, particularly inside regions of delayed cooling.

6.3. Stellar Properties

In this section, we carry out a global analysis of the stellar components in the CosmoRun simulations, but only focusing on their spatial distribution and metallicity. A more detailed analysis of the stellar components, including the kinematics, SFHs, in situ versus ex situ origin, and low-z evolution, will be presented in a future paper by the AGORA Collaboration.

In Section 5.4.2 for Cal-4, we have examined the stellar mass growth histories (Figure 12). There, we detect occasional increases in stellar masses in most codes—the kinds of increases that are not contemporaneous between the codes. In fact, these are signs of major mergers, which can be best observed in the stellar surface density maps in Figure 21. The mean virial radius, R_{200}, at each redshift (see the Figure 15 caption) is marked with a white dashed circle in each panel. In this figure, it is easier to perceive that major/minor mergers do not occur at the same time in every simulation due to the aforementioned timing discrepancy (see Sections 5.4.2 and 6.2). The $z = 4$ row is particularly interesting. By $z = 4$, most codes have gone through a recent major merger event, but they are at different stages of halo relaxation. This observation reminds us of the need to be careful when comparing properties of galaxy-scale systems in cosmological simulations between different codes; it is indeed prudent to avoid times when a strong perturbation is ongoing. The simulations presented here will be further analyzed in a future paper, also at lower redshifts, when major mergers are rare and comparisons are more straightforward.

We conclude this subsection by investigating stellar metallicities and comparing the results with the distribution of metals in the gas component. By construction, stars form in regions where gas reaches the imposed star formation threshold; thus they inherit the properties of their progenitor gas. Among the inherited properties, metallicity is the one that should follow a similar trend between stars and the high-density gas. Additionally, in the gas metallicity PDF within R_{200} (Figures 19 and 20), we expect to find that a significant fraction of gas in the high-density, high-metallicity bins is star-forming. This argument is in agreement with what we observe in Figure 22, in which we show the stellar mass per metallicity bin. As can be also inferred from Figures 17 and 19, the stellar metallicity distribution peaks at a similar value to the gas metallicity in each code. Nevertheless, the distribution tends to be narrower in the stellar metallicities (Figure 22) than in the gas metallicities (Figure 19), as most star particles form in the densest pockets of gas. The low-metallicity stars could be either the early generation of stars formed in the gas that have not been heavily metal-enriched yet, or the later generation of stars formed in the CGM only lightly metal-enriched by galactic outflows.

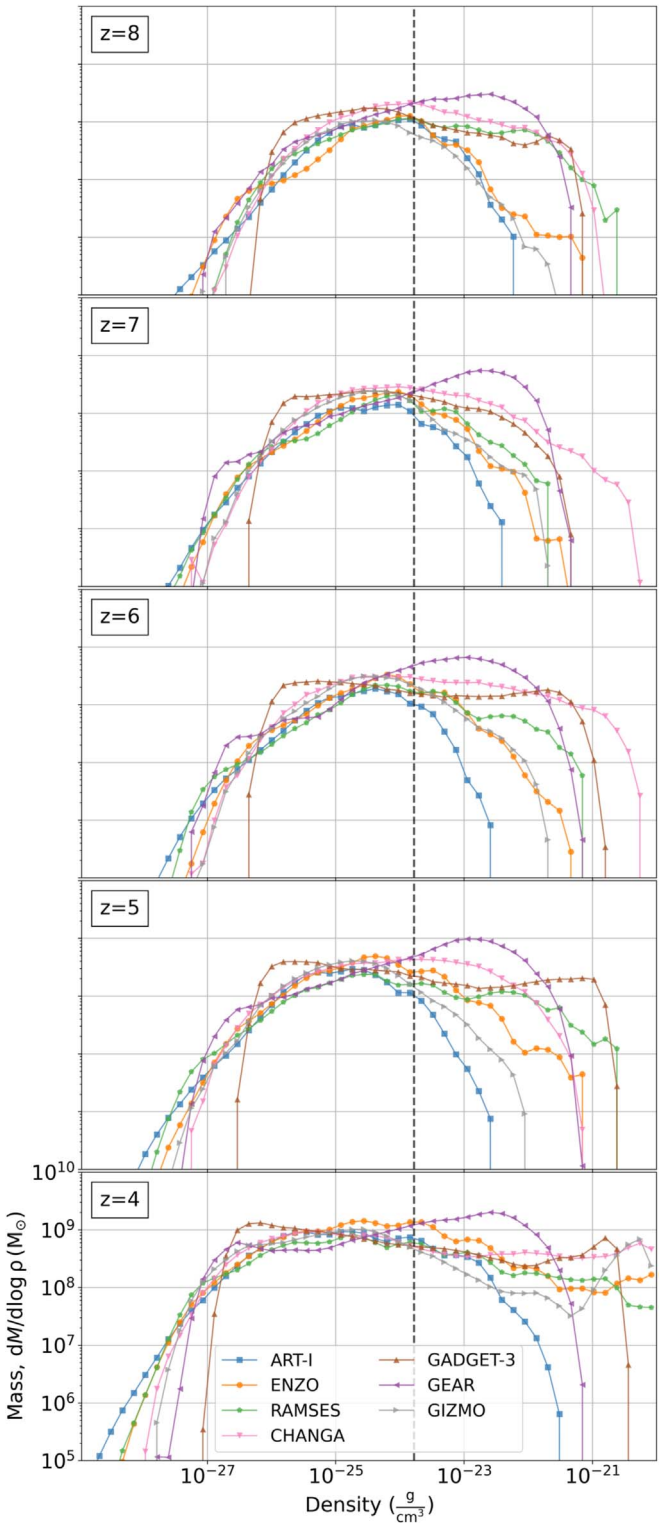

Figure 17. Distribution of gas mass as a function of gas density at $z = 8$, 7, 6, 5, and 4 from our CosmoRun simulation suite. Each panel is for all the gas inside the target progenitor's R_{200}. The vertical black dashed line denotes the star formation threshold, $n_{\rm H,thres} = 1\ {\rm cm}^{-3}$. See Section 6.2 for more information on CosmoRun and this figure.

6.4. CGM Properties

The AGORA Collaboration plans to work on a full analysis of the CGM properties and evolution of the presented CosmoRun simulations, from high z down to $z = 2$. The results of this extensive analysis will be presented in a forthcoming paper. In this section, however, we demonstrate how multiplatform studies like AGORA could be useful to better understanding the thermal and kinematic states of the CGM, in which disparities exist between contemporary cosmological simulations carried out with different codes, by presenting the first analysis of gas kinematics in four different temperature bins at $z = 4$. The temperature bins are defined following the observationally motivated temperature thresholds proposed in Roca-Fàbrega et al. (2019) and in Strawn et al. (2021).

In Figure 23, we show the probability distributions of the velocity magnitude (top row) and the radial velocity (bottom row) for the gas inside a sphere of radius R_{200} from the center of the target progenitor galaxy. The panels are for all the gas, cold gas ($T < 10^{3.8}$ K), cool gas ($10^{3.8} < T < 10^{4.5}$ K), warm gas ($10^{4.5} < T < 10^{6.5}$ K), and hot gas ($T > 10^{6.5}$ K) from left to right. The velocity magnitude PDFs (top row) show that there is reasonably good agreement on the kinematics of the gas. This agreement is particularly good in the cool and warm gas; in these temperature phases, the mesh-based codes and the particle-based codes agree well with each other. The convergence is not as good in the hot gas, though, where ART-I and ENZO exhibit slightly larger gas fractions with high velocity than the rest of the participating codes, due to stronger feedback-driven outflows (rightmost panel; as discussed in Section 6.2). The RAMSES run presented here shows lower velocities than ART-I and ENZO in the hot gas component as expected from our analysis of metal distribution (see a full discussion in Section 6.2). Additionally, in the CHANGA, GADGET-3, GEAR, and GIZMO runs, the hot gas with the highest velocities typically belongs to regions with very low density, which are not well represented by their particle-based approach. In agreement with our conclusions on the gas metallicity distribution (see Section 6.2.2), GEAR generates the slowest outflows, keeping most of the metals in the dense gas around the galaxy.

In the bottom row of Figure 23, we show the distribution of gas mass in radial velocity bins. Radial velocity informs us of the presence of inflowing or outflowing gas, and the strength thereof. As discussed in the previous paragraph and in Section 6.2, the strong feedback-driven outflows in ART-I and ENZO are evident in the hot gas phase (rightmost panel; also in the warm phase for ART-I). This outflowing hot gas transports a large fraction of metals to the IGM, leaving the CGM in ART-I and ENZO with lower metallicity relative to the other codes. The RAMSES, CHANGA, GADGET-3, GIZMO, and particularly GEAR runs do not show as strong outflows as those in ART-I or ENZO, keeping most of the metals and gas inside the CGM (as also seen in Figures 17 and 18). The cool gas follows a smooth distribution centered at zero velocity but is slightly inflowing (third panel from the left), with very good agreement among all the codes.

The very preliminary analysis of gas properties in the CGM and, in particular, of its kinematics in four different temperature bins teaches us that the kinematics of cold and hot gas is a good tracer of differences in the adopted stellar feedback prescriptions. We suggest that research groups interested in testing their feedback models include the study of cold and hot gas kinematics in their comparisons.

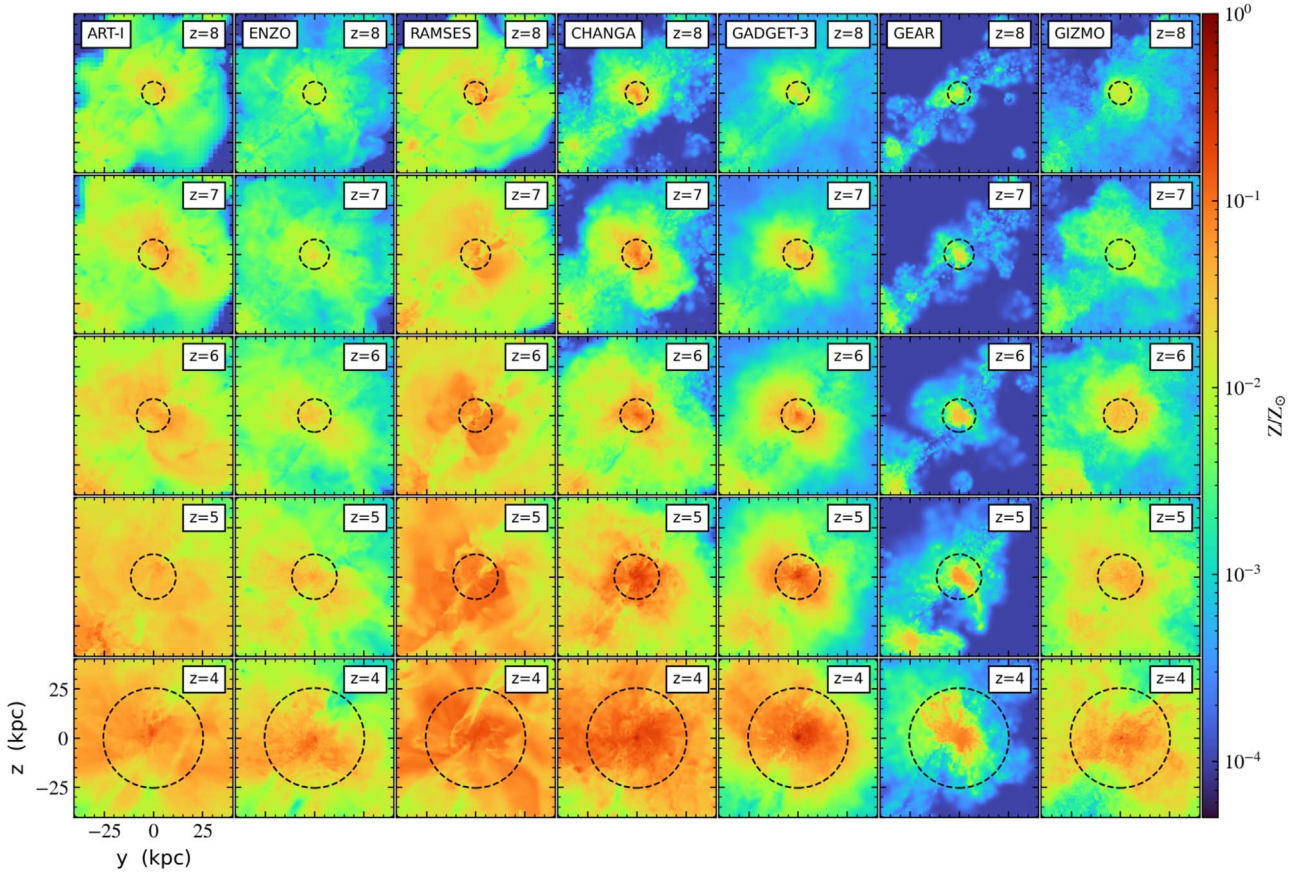

Figure 18. Same as Figures 15 and 16, but now showing density-square-weighted projections of gas metallicity in our CosmoRun simulation suite. Colors represent the metallicity in units of $Z_{\odot}$. The units are proper kiloparsecs. See Section 6.2 for more information on CosmoRun and this figure.

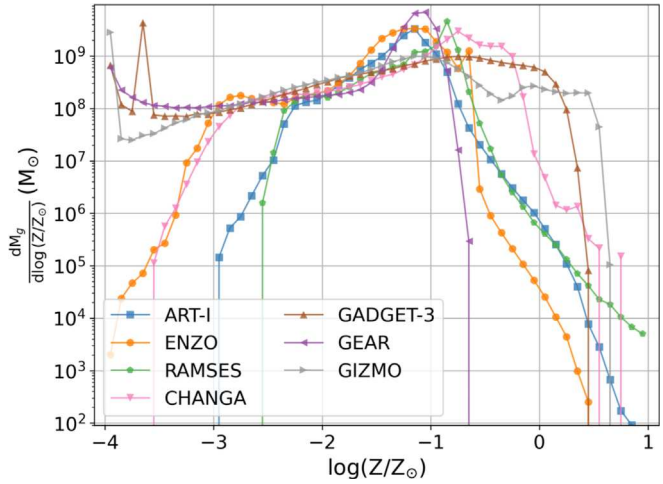

Figure 19. Distribution of gas mass as a function of gas metallicity at $z = 4$ for all the gas inside the target progenitor's R_{200} in our CosmoRun simulation suite. The y-axis range is kept identical to that in Figure 22 for easier comparison. See Section 6.2 for more information on CosmoRun and this figure.

7. Discussion and Conclusion

In this paper, we have presented a suite of seven high-resolution cosmological zoom-in simulations to $z = 4$ of a halo with a Milky Way mass at $z = 0$, obtained using seven contemporary astrophysical simulation codes—three AMR codes and four SPH codes—widely used in numerical galaxy formation. The physics prescriptions in the simulations include the common gas cooling and heating scheme by Grackle v3.1.1, which is similar to what was used in the previous AGORA comparisons, and standardized AGORA subgrid physics, such as star formation and stellar evolution (Section 3.1). However, the code groups participating in the comparison use a stellar feedback prescription that resembles the one most widely used in their code community for research (Section 3.2). The simulations also account for the effects of cosmological processes such as the expansion of the universe and the cosmic UVB radiation emitted by massive stars and quasars.

The simulations presented here have been obtained after a careful, four-step process of calibrations (Section 5). The calibration strategy designed by the collaboration is to reduce the number of tunable simulation parameters to be accounted for when studying the effects of stellar feedback on galaxy evolution. By completing this set of calibrations, the participating code groups establish a common ground on which to make a robust and unbiased comparison of different simulations focusing on stellar feedback effects on the gas and SFH of the target galaxy. The calibration procedure includes four steps. In the first step (Cal-1) the code groups control the effects of different gravity and hydrodynamics solvers and refinement strategies in radiative cooling/heating-free simulations. In the second step (Cal-2), we ensure that the Grackle cooling and UVB are correctly implemented in each code. The third step (Cal-3) aims for convergence in the total stellar mass

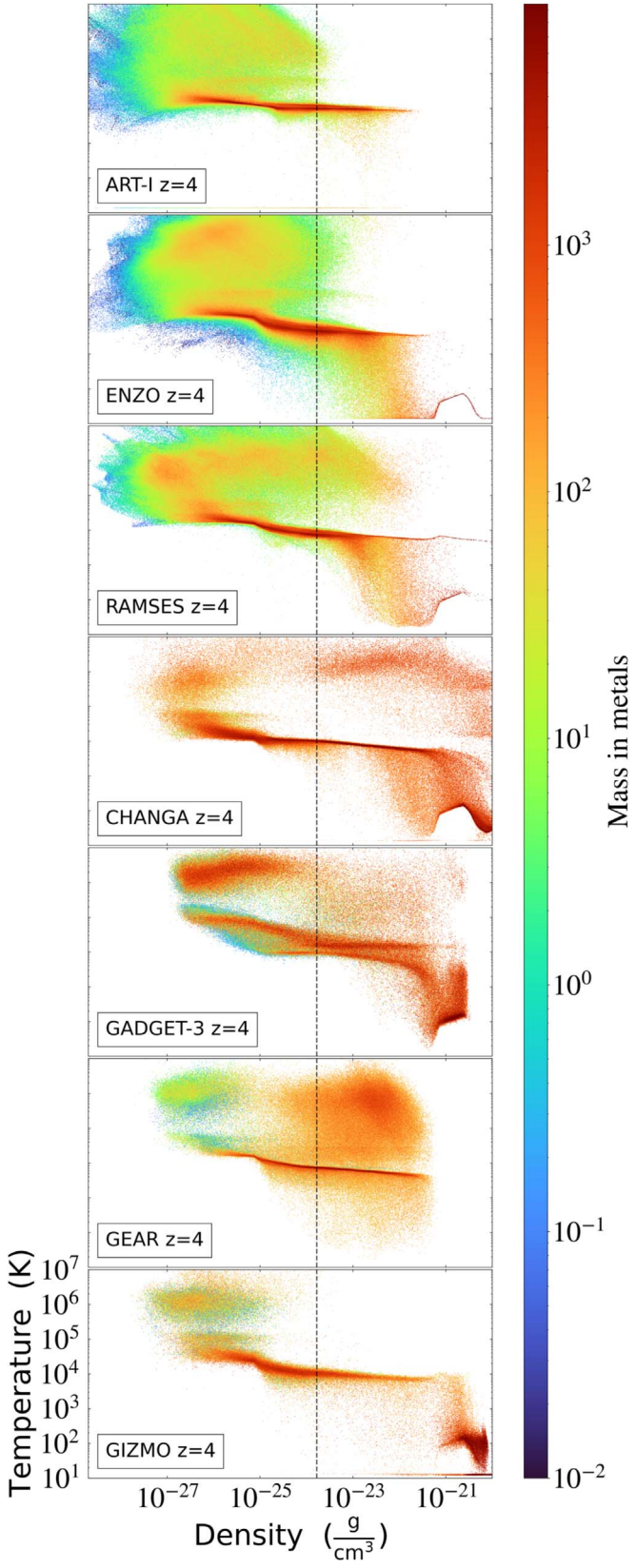

Figure 20. Same as Figure 14, but now with colors representing the total metal mass in each two-dimensional bin in our CosmoRun simulation suite. Note that the PDF is for the gas within R_{200} of the center of the target galaxy in the CosmoRun simulations. A sphere of radius R_{200} encloses the main galaxy and CGM, but not the IGM. See Section 6.2 for more information on CosmoRun and this figure.

produced with the common star formation prescription in stellar-feedback-free simulations. In the last calibration step (Cal-4), we ask each code group to test a stellar feedback prescription that is as close to the most commonly used one in each code community as possible, while aiming for convergence of the stellar-to-halo mass ratio at $z = 4$ to the prediction by semiempirical models. Designing and executing the calibration procedure has required formidable efforts by collaboration members to (re)run the simulations while revising, when necessary, the physical prescriptions they use for the final cosmological simulations.

After all the participating code groups successfully complete the calibration steps, we achieve a suite of cosmological zoom-in simulations with very similar mass assembly histories down to $z = 4$ (CosmoRun; Section 6). With numerical accuracy that resolves the internal structure of a target halo ($\lesssim$100 physical pc at $z = 4$), we find that the codes overall agree well with one another in many aspects. We argue that, if adequately tested in accordance with our proposed calibration steps and common parameters, modern high-resolution cosmological zoom-in simulations produce robust results and their predictive power can be maximized. While this calibration does lead to substantial agreement on critical parameters, differences still remain between the codes—in the properties of the gas, stars, and CGM—due to the different stellar feedback strategy adopted in each of the participating codes, as well as the diversity in the implementation of hydrodynamics. We show that the gas distribution in the density–temperature space is globally affected by differences in the stellar feedback, particularly in the coldest and hottest gas, while achieving solid convergence in the cool and warm gas. We also confirm that the spatial distribution of gas metallicity from metals released in the SN explosion is a key parameter when testing stellar feedback prescriptions in cosmological models. This is because gas metallicity plays an important role in the gas cooling rates, amplifying the differences in the feedback prescriptions. A similar effect is observed when analyzing stellar metallicities. We also confirm that the expected timing discrepancies in halo mergers need to be accounted for when making code-to-code comparisons, since variations in the host's post-merger relaxation highly impact the gas properties. The analysis presented in this paper, which includes only five redshift epochs (i.e., $z = 8$, 7, 6, 5, and 4), serves as a first presentation of our suite of cosmological zoom-in simulations, and we are currently running them down to lower redshift and saving snapshots at finer timesteps.

It is important to briefly note a few points about our study presented in this work: (1) Our comparison in this paper across different code platforms is possible only because we have established a solid baseline through rigorous calibration steps (Section 5). The proposed calibration procedure has enabled us to trust that any differences can only be attributed to the chosen stellar feedback prescriptions and the (relatively minor) intrinsic variations of the codes' numerics. (2) The process of running cosmological simulations through multiple calibration steps and production stages has required herculean effort from many AGORA members. It has also been facilitated by close discussions between the code representatives, through three workshops and more than 30 teleconferences (for the CosmoRun simulations alone; as of 2021 May) hosted by the collaboration. This type of interplatform collaboration is somewhat novel in the field of numerical cosmology. (3) Throughout

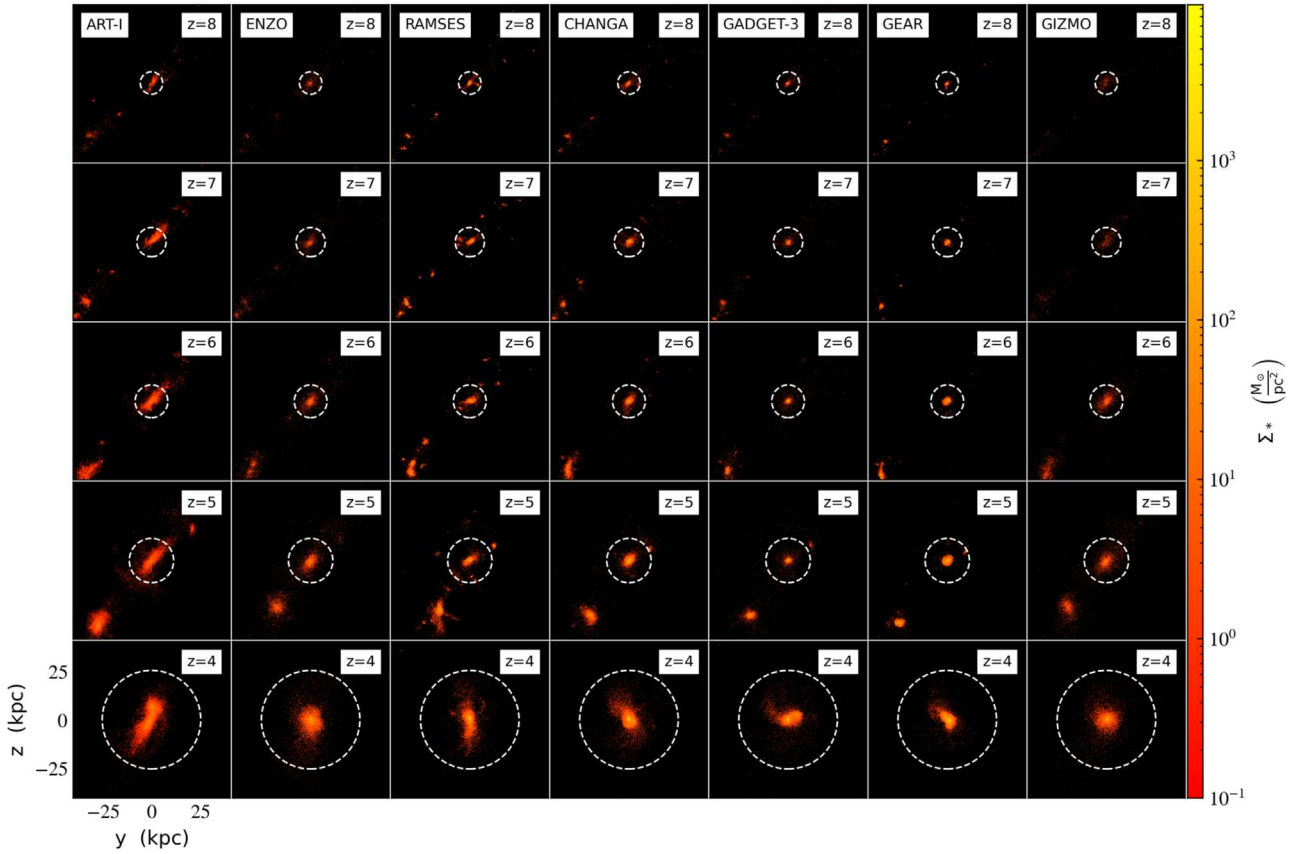

Figure 21. Same as Figures 15, 16, and 18, but now presenting stellar surface densities from our CosmoRun simulation suite. Colors represent the total stellar mass in each two-dimensional bin. The units are proper kiloparsecs. See Section 6.3 for more information on CosmoRun and this figure.

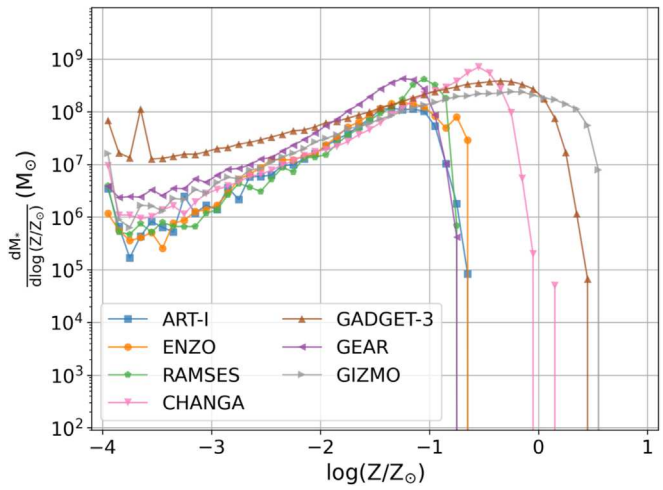

Figure 22. Distribution of stellar mass as a function of stellar metallicity at $z = 4$ for all the stars inside the target progenitor's R_{200} in our CosmoRun simulation suite. The y-axis range is kept identical to that in Figure 19 for easier comparison. See Section 6.3 for more information on CosmoRun and this figure.

this invaluable learning process, participants have used AGORA as a forum in which to talk to and learn from one another about other codes, and sometimes surprisingly, about their own. Many participants have been able to improve their codes and simulation strategies. The new versions of GRACKLE and yt were tested on multiple code platforms during this work, providing useful feedback to the respective developer communities.

We pride ourselves on our contribution to the galaxy formation community, of helping to maintain the reproducibility of galaxy formation simulations in general. AGORA helps to raise the predictive power of numerical experiments—this time, in particular, that of cosmological zoom-in simulations—in building and testing the theory of structure formation in the universe, thereby benefiting researchers who rely on the robustness of simulations. Furthermore, we have demonstrated how a multiplatform approach like AGORA's could be useful to better understanding how the universe works. For example, in AGORA, the thermal and kinematic states of the CGM—in which disparities exist between contemporary numerical simulations on different code platforms—can be easily investigated with multiple codes and increased fidelity, as showcased in Section 6. Indeed, AGORA enables a well-controlled science case in which we can test various stellar feedback prescriptions and confront simulations with those from other codes. The novel infrastructure presented in this work will provide the AGORA community (or the broader simulation community) with a tool with which to undertake a number of new comparison projects, including the analysis of CGM properties in simulations with different stellar feedbacks, the formation of clumps at high redshift, and many others. It should be noted that the code groups involved in other ongoing projects in AGORA or in any upcoming new projects are not limited to the seven code groups that participated in this paper. Our collaboration is open to the participation of new code groups, and we encourage interested community members to test their code's compatibility on their own, by adopting the common initial conditions, the common physics package, and

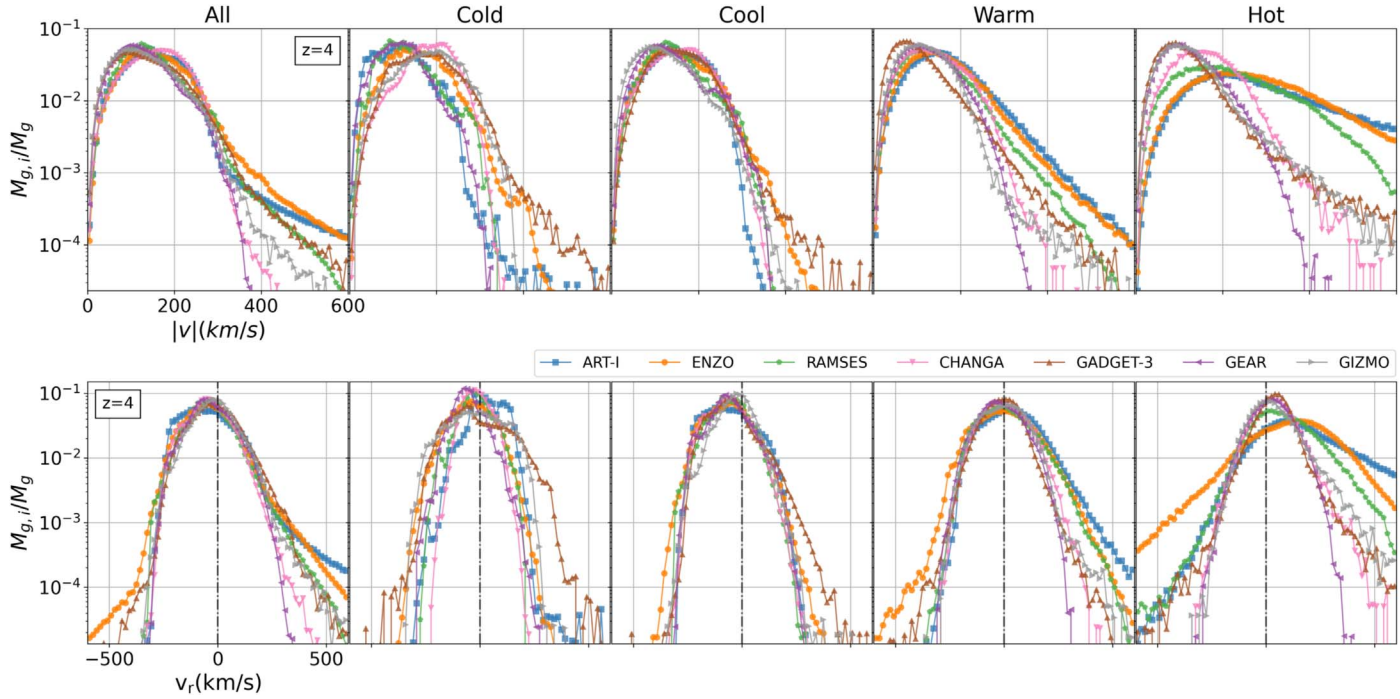

Figure 23. Distribution of gas mass as a function of velocity at $z = 4$—velocity magnitude (top) and radial velocity (bottom)—for the gas inside the target progenitor's R_{200} in our CosmoRun simulation suite. The y-axis indicates the fraction of gas mass in each velocity bin with respect to the total mass in each temperature phase. The panels are for all the gas, cold gas ($T < 10^{3.8}$ K), cool gas ($10^{3.8} < T < 10^{4.5}$ K), warm gas ($10^{4.5} < T < 10^{6.5}$ K), and hot gas ($T > 10^{6.5}$ K) from left to right. See Section 6.4 for more information on CosmoRun and this figure.

the proposed calibration steps, and comparing their results with the ones from the models presented by the AGORA Collaboration.

We thank all of our colleagues who participate in the AGORA Project for their collaborative spirit, which has allowed the AGORA Collaboration to remain strong as a platform to foster and launch multiple science-oriented comparison efforts. We thank Aldo Rodríguez-Puebla for sharing results from abundance matching semiempirical models, and Volker Springel for providing the original versions of GADGET-3 to be used in the AGORA Project. We also thank the anonymous referee for insightful comments and suggestions. This research used resources of the National Energy Research Scientific Computing Center, a user facility supported by the Office of Science of the U.S. Department of Energy under contract No. DE-AC02-05CH11231. S.R.-F. acknowledges support from a Spanish postdoctoral fellowship, under grant No. 2017-T2/TIC-5592. His work has been supported by the Madrid Government (Comunidad de Madrid–Spain) under the Multiannual Agreement with Complutense University in the line Program to Stimulate Research for Young Doctors in the context of V PRICIT (Regional Programme of Research and Technological Innovation). He also acknowledges financial support from the Spanish Ministry of Economy and Competitiveness under grant Nos. AYA2016-75808-R, AYA2017-90589-REDT, and S2018/NMT-429, and from CAM-UCM under grant No. PR65/19-22462. J.K. acknowledges support from the Samsung Science and Technology Foundation under project No. SSTF-BA1802-04. His work was also supported by the National Institute of Supercomputing and Networking, Korea Institute of Science and Technology Information, with supercomputing resources, including technical support (grants KSC-2018-CRE-0052 and KSC-2019-CRE-0163). K.N. acknowledges support from MEXT/JSPS KAKENHI grant Nos. JP17H01111, 19H05810, and 20H00180, as well as travel support from Kavli IPMU, World Premier Research Center Initiative, where part of this work was conducted. A.L. acknowledges funding by the MIUR under the grant PRIN 2017-MB8AEZ. D.C. is a Ramon Cajal Researcher and is supported by Ministerio de Ciencia, Innovación y Universidades (FEDER) under research grant PGC2018-094975-C21. H.V. acknowledges support from PAPIIT of Universidad Nacional Autónoma de México (UNAM) under grant No. IN101918 and also from Centro Nacional de Supercomputo (CNS-IPICYT-CONACYT). The ART-I simulations were performed on the BRIGIT/EOLO cluster at Centro de Proceso de Datos, Universidad Complutense de Madrid, and on the STÓCATL supercomputer at Instituto de Astronomía de la UNAM. The RAMSES simulations were performed on the MIZTLI supercomputer at LANCAD, UNAM, within the research project LANCAD-UNAM-DGTIC-151 and on Laboratorio Nacional de Supercómputo del Sureste, CONACYT. The CHANGA simulations were performed on the ATÓCATL supercomputer at Instituto de Astronomía de la UNAM, and on the Extreme Science and Engineering Discovery Environment (XSEDE) allocations TG-AST20020 and TG-MCA94P018. XSEDE is supported by the National Science Foundation grant ACI-1053575. The GADGET3-OSAKA simulations and analyses were performed on the XC50 systems at the Center for Computational Astrophysics of the National Astronomical Observatory of Japan, on OCTOPUS at the Cybermedia Center of Osaka University, and on Oakforest-PACS at the University of Tokyo as part of the HPCI System Research Project (hp190050 and hp200041). The publicly available ENZO and yt codes used in this work are the products of collaborative efforts by many independent scientists from numerous institutions around the world. Their commitment to open science has helped make this work possible.

ORCID iDs

Santi Roca-Fàbrega https://orcid.org/0000-0002-6299-152X
Ji-hoon Kim https://orcid.org/0000-0003-4464-1160
Kentaro Nagamine https://orcid.org/0000-0001-7457-8487
Alessandro Lupi https://orcid.org/0000-0001-6106-7821
Daniel Ceverino https://orcid.org/0000-0002-8680-248X
Joel R. Primack https://orcid.org/0000-0001-5091-5098
Tom Abel https://orcid.org/0000-0002-5969-1251
Avishai Dekel https://orcid.org/0000-0003-4174-0374
Oliver Hahn https://orcid.org/0000-0001-9440-1152
Cameron Hummels https://orcid.org/0000-0002-3817-8133
Romain Teyssier https://orcid.org/0000-0001-7689-0933
Matthew J. Turk https://orcid.org/0000-0002-5294-0198

References

Aoyama, S., Hou, K.-C., Hirashita, H., Nagamine, K., & Shimizu, I. 2018, MNRAS, 478, 4905
Aoyama, S., Hou, K.-C., Shimizu, I., et al. 2017, MNRAS, 466, 105
Brummel-Smith, C., Bryan, G., Butsky, I., et al. 2019, JOSS, 4, 1636
Bryan, G. L., Norman, M. L., O'Shea, B. W., et al. 2014, ApJS, 211, 19
Ceverino, D., Dekel, A., & Bournaud, F. 2010, MNRAS, 404, 2151
Ceverino, D., & Klypin, A. 2009, ApJ, 695, 292
Ceverino, D., Klypin, A., Klimek, E. S., et al. 2014, MNRAS, 442, 1545
Ceverino, D., Primack, J., Dekel, A., & Kassin, S. A. 2017, MNRAS, 467, 2664
Dehnen, W., & Aly, H. 2012, MNRAS, 425, 1068
Dubois, Y., Volonteri, M., Silk, J., et al. 2015, MNRAS, 452, 1502
Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, RMxAA, 49, 137
Grisdale, K., Agertz, O., Renaud, F., et al. 2019, MNRAS, 486, 5482
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Hahn, O., & Abel, T. 2011, MNRAS, 415, 2101
Hayward, C. C., & Hopkins, P. F. 2017, MNRAS, 465, 1682
Hernquist, L., & Katz, N. 1989, ApJS, 70, 419
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hopkins, P. F. 2013, MNRAS, 428, 2840
Hopkins, P. F. 2015, MNRAS, 450, 53
Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 477, 1578
Hosono, N., Saitoh, T. R., & Makino, J. 2016, ApJS, 224, 32
Hubber, D. A., Falle, S. A. E. G., & Goodwin, S. P. 2013, MNRAS, 432, 711
Jia, J.-Y., Zhu, W.-S., Gao, L., & Feng, L.-L. 2020, RAA, 20, 095
Kacprzak, G. G., Pointon, S. K., Nielsen, N. M., et al. 2019, ApJ, 886, 91
Keller, B. W., Wadsley, J., Benincasa, S. M., & Couchman, H. M. P. 2014, MNRAS, 442, 3013
Kim, J.-h., Abel, T., Agertz, O., et al. 2014, ApJS, 210, 14
Kim, J.-H., Agertz, O., Teyssier, R., et al. 2016, ApJ, 833, 202
Kim, J.-h., Wise, J. H., Abel, T., et al. 2019, ApJ, 887, 120
Kobayashi, C., Tsujimoto, T., & Nomoto, K. 2000, ApJ, 539, 26
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
Kravtsov, A. V. 2003, ApJL, 590, L1
Kravtsov, A. V., Klypin, A. A., & Khokhlov, A. M. 1997, ApJS, 111, 73
Lehner, N., Berek, S. C., Howk, J. C., et al. 2020, ApJ, 900, 9
Lukić, Z., Stark, C. W., Nugent, P., et al. 2014, MNRAS, 446, 3697
Lupi, A., Pallottini, A., Ferrara, A., et al. 2020, MNRAS, 496, 5160
Matteucci, F., & Greggio, L. 1986, A&A, 154, 279
Morris, J. P. 1996, PASA, 13, 97
Nuñez-Castiñeyra, A., Nezri, E., Devriendt, J., & Teyssier, R. 2021, MNRAS, 501, 62
Oh, B. K., Smith, B. D., Peacock, J. A., & Khochfar, S. 2020, MNRAS, 497, 5203
Okamoto, T., Eke, V. R., Frenk, C. S., & Jenkins, A. 2005, MNRAS, 363, 1299
Power, C., Navarro, J. F., Jenkins, A., et al. 2003, MNRAS, 338, 14
Raiteri, C. M., Villata, M., & Navarro, J. F. 1996, A&A, 315, 105
Revaz, Y., Arnaudon, A., Nichols, M., Bonvin, V., & Jablonka, P. 2016, A&A, 588, A21
Revaz, Y., & Jablonka, P. 2012, A&A, 538, A82
Revaz, Y., & Jablonka, P. 2018, A&A, 616, A96
Roca-Fàbrega, S., Dekel, A., Faerman, Y., et al. 2019, MNRAS, 484, 3625
Roca-Fàbrega, S., Kim, J.-h., Primack, J. R., et al. 2020, arXiv:2001.04354
Rodríguez-Puebla, A., Primack, J. R., Avila-Reese, V., & Faber, S. M. 2017, MNRAS, 470, 651
Rosdahl, J., Schaye, J., Dubois, Y., Kimm, T., & Teyssier, R. 2017, MNRAS, 466, 11
Saitoh, T. R. 2017, AJ, 153, 85
Saitoh, T. R., & Makino, J. 2009, ApJL, 697, L99
Saitoh, T. R., & Makino, J. 2013, ApJ, 768, 44
Scannapieco, C., Wadepuhl, M., Parry, O. H., et al. 2012, MNRAS, 423, 1726
Shen, S., Wadsley, J., & Stinson, G. 2010, MNRAS, 407, 1581
Shimizu, I., Todoroki, K., Yajima, H., & Nagamine, K. 2019, MNRAS, 484, 2632
Shin, E.-J., Jung, M., Kwon, G., et al. 2020, ApJ, 899, 25
Smith, B. D., Bryan, G. L., Glover, S. C. O., et al. 2017, MNRAS, 466, 2217
Springel, V. 2005, MNRAS, 364, 1105
Springel, V. 2010, MNRAS, 401, 791
Stinson, G., Seth, A., Katz, N., et al. 2006, MNRAS, 373, 1074
Strawn, C., Roca-Fàbrega, S., Mandelker, N., et al. 2021, MNRAS, 501, 4948
Suresh, J., Bird, S., Vogelsberger, M., et al. 2015, MNRAS, 448, 895
Taylor, P. A., & Miller, J. C. 2012, MNRAS, 426, 1687
Teyssier, R. 2002, A&A, 385, 337
Tornatore, L., Borgani, S., Dolag, K., & Matteucci, F. 2007, MNRAS, 382, 1050
Tsujimoto, T., Nomoto, K., Yoshii, Y., et al. 1995, MNRAS, 277, 945
Turk, M. J., Smith, B. D., Oishi, J. S., et al. 2011, ApJS, 192, 9
Wadsley, J. W., Keller, B. W., & Quinn, T. R. 2017, MNRAS, 471, 2357
Wiersma, R. P. C., Schaye, J., Theuns, T., Vecchia, C. D., & Tornatore, L. 2009, MNRAS, 399, 574