Code Comparison in Galaxy-scale Simulations with Resolved Supernova Feedback: Lagrangian versus Eulerian Methods

Chia-Yu Hu (胡家瑜)1, Matthew C. Smith2,3, Romain Teyssier4, Greg L. Bryan5,6, Robbert Verbeke7, Andrew Emerick5, Rachel S. Somerville5, BlakesleyBurkhart5,8, Yuan Li (黎原)9, John C. Forbes5, and Tjitske Starkenburg10

1 Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse 1, D-85748 Garching, Germany; cyhu.astro@gmail.com
2 Universität Heidelberg, Zentrum für Astronomie, Institut für theoretische Astrophysik, Albert-Ueberle-Str. 2, D-69120 Heidelberg, Germany
3 Max-Planck-Institut für Astromie, Königstuhl 17, D-69117 Heidelberg, Germany
4 Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
5 Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA
6 Department of Astronomy, Columbia University, Pupin Physics Laboratories, New York, NY 10027, USA
7 Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
8 Department of Physics and Astronomy, Rutgers University, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA
9 Department of Physics, University of North Texas, Denton, TX 76203, USA
10 Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, Northwestern University, 1800 Sherman Avenue, Evanston, IL 60201, USA

Received 2022 August 22; revised 2023 April 14; accepted 2023 April 21; published 2023 June 16

Abstract

We present a suite of high-resolution simulations of an isolated dwarf galaxy using four different hydrodynamical codes: GIZMO, AREPO, GADGET, and RAMSES. All codes adopt the same physical model, which includes radiative cooling, photoelectric heating, star formation, and supernova (SN) feedback. Individual SN explosions are directly resolved without resorting to subgrid models, eliminating one of the major uncertainties in cosmological simulations. We find reasonable agreement on the time-averaged star formation rates as well as the joint density–temperature distributions between all codes. However, the Lagrangian codes show significantly burstier star formation, larger SN-driven bubbles, and stronger galactic outflows compared to the Eulerian code. This is caused by the behavior in the dense, collapsing gas clouds when the Jeans length becomes unresolved: Gas in Lagrangian codes collapses to much higher densities than that in Eulerian codes, as the latter is stabilized by the minimal cell size. Therefore, more of the gas cloud is converted to stars and SNe are much more clustered in the Lagrangian models, amplifying their dynamical impact. The differences between Lagrangian and Eulerian codes can be reduced by adopting a higher star formation efficiency in Eulerian codes, which significantly enhances SN clustering in the latter. Adopting a zero SN delay time reduces burstiness in all codes, resulting in vanishing outflows as SN clustering is suppressed.

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

1. Introduction

Tremendous progress has been made in hydrodynamical simulation of galaxy formation in the last decade (see Somerville & Davé 2015 and references herein). Starting with initial conditions from the cosmic microwave background, the simulated galaxies evolving across the cosmic time are remarkably realistic in many aspects, such as their masses, sizes, metallicities, colors, morphologies, etc. These simulations provide crucial information that is often observationally inaccessible and therefore have routinely been used to study the physical processes that shape the observed galaxies.

However, galaxy formation is fundamentally a multiscale problem that spans several orders of magnitude both spatially and temporally. The computational cost of simulating the relevant dynamical range is impractically high and will remain so in the foreseeable future. Therefore, phenomenological subgrid models are required to account for the physical processes on unresolved scales with several tunable parameters, rendering their predictive power somewhat ambiguous (see Naab & Ostriker 2017 and references therein).

On the other hand, small-scale simulations have provided valuable insights from a “bottom-up” perspective (Walch et al. 2015; Girichidis et al. 2016; Ibáñez-Mejía et al. 2016; Gatto et al. 2017; Li et al. 2017; Kim & Ostriker 2017, 2018; Kim et al. 2020; Hu et al. 2021). These simulations focus on a patch of the interstellar medium (ISM) on kiloparsec scales and therefore can achieve much higher resolution, following the turbulence and gravitational collapse that shapes the multiphase ISM down to much smaller scales and higher densities. More importantly, feedback from supernova (SN) explosions can be faithfully modeled even when the energy-conserving Sedov–Taylor phase is resolved, which requires a spatial resolution of ~4 pc in Eulerian codes (Kim & Ostriker 2015; Simpson et al. 2015) or a mass resolution of ~10M⊙in Lagrangian codes (Hu 2019; Steinwandl et al. 2020) under typical ISM conditions. The difficulty in resolving SN feedback in cosmological simulations is a long-standing problem and has motivated a plethora of subgrid models. Dramatically different simulation results arise primarily from the differences in the adopted feedback prescriptions while the differences in the hydrodynamical techniques only play a secondary role (e.g., Scannapieco et al. 2012). Resolving SN feedback therefore eliminates a major source of uncertainty.
However, the periodic boundary conditions adopted in ISM-patch simulations place a fundamental limitation on their applicability on large scales such as the lack of spiral waves (Smith et al. 2020) and potentially unrealistic behavior in galactic outflows (Martizzi et al. 2016).

Recently, galaxy-scale simulations that can resolve individual SN feedback in an isolated galaxy have become feasible (Forbes et al. 2016; Hu et al. 2016, 2017; Smith et al. 2018, 2021; Hu 2019; Emerick et al. 2019; Gutcke et al. 2021; Smith 2021; Hislop et al. 2022). They have more realistic boundary conditions than the ISM-patch simulations and can accurately capture the large-scale properties. Dwarf galaxies, thanks to their small sizes, are the primary targets for this type of simulation. More ambitious large-scale properties. Dwarf galaxies, thanks to their small sizes, are therefore subject to the caveat of the neglected physics, complexity that can capture the essential physics. Our results are therefore subject to the caveat of the neglected physics, which we will discuss in detail in Section 4.

This paper is organized as follows. In Section 2, we introduce our numerical framework and the hydrodynamical codes. In Section 3, we present our simulation results, demonstrating the striking differences between Lagrangian and Eulerian codes in the burstiness of star formation, gas morphology, and galactic outflows as a result of SN clustering. In Section 4, we discuss the implications and limitations of our results and compare them with previous work. In Section 5, we summarize our results.

2. Numerical Methods

2.1. Hydrodynamical Codes

We use four different codes for gravity and hydrodynamics: one Eulerian code (RAMSES), one moving mesh code (AREPO), and two Lagrangian codes (GIZMO and GADGET-3), which we briefly describe as follows.

GADGET-3 (hereafter GADGET for short) (Springel 2005) is a particle-based code. Gravity is solved using a “tree code” (Barnes & Hut 1986) where short-range forces are calculated pairwise while long-range forces are approximated by the center of mass of a tree node. Since it is designed for collisionless N-body problems, gravitational forces at small distances are reduced by softening, and different softening lengths can be used for different components, such as gas, stars, and dark matter. Hydrodynamics is solved by the smoothed particle hydrodynamics (SPH) method (Lucy 1977). We adopt the SPH GAL implementation in Hu et al. (2014) that includes several improvements over traditional SPH methods; these improvements include the pressure-energy formulation (Hopkins 2015), the Wendland kernel (Dehnen & Aly 2012), variable artificial viscosity and conduction (Price 2008; Cullen & Dehnen 2010), and the time-step limiter (Durier & Dalla Vecchia 2012). Both gravity and hydrodynamics solvers are Lagrangian in the sense that particles follow the motions of the gas, stars, and dark matter. The initial mass carried by each particle is conserved as there are no mass fluxes between particles.

GIZMO (Hopkins 2015) is a multimethod code built on the code GADGET, featuring the meshless Godunov method (Gaburov & Nitadori 2011) for hydrodynamics. We adopt the meshless finite-mass (FMF) method proposed by Hopkins (2015) in this work. In FMF, the simulation domain is divided into overlapping cells defined by the particle distribution and a kernel function with a finite support radius. The Riemann problem is solved at the interfaces between cells. The cells move with the gas so that the mass fluxes between cells are assumed to be zero. GIZMO adopts the same gravity solver as GADGET and is a Lagrangian code when the FMF solver is used.

AREPO (Springel 2010; Pakmor et al. 2016; Weinberger et al. 2020) uses a second-order accurate finite-volume scheme on an unstructured, moving mesh. The simulation domain is discretized by a Voronoi tessellation used in contrast to GADGET and GIZMO. In the standard usage of the code, adopted here, a refinement (de-refinement) scheme is used to split (merge) cells in order to enforce a constant mass resolution within a factor of 2. AREPO adopts a gravity solver similar to those employed in GADGET and GIZMO.

RAMSES (Teyssier 2002) is an Eulerian code with octree-based adaptive mesh refinement (AMR). The N-body solver is based on the adaptive-particle-mesh method. The Poisson equation is solved using the multigrid method with Dirichlet boundary conditions at level boundaries (Guilet & Teyssier 2011). Hydrodynamics is solved using the MUSCL scheme, a second-order finite-volume Godunov method, and the HLLC Riemann solver. For RAMSES, we used here a quasi-Lagrangian refinement criterion based on a target mass common to all codes in this work. We use a box size of 450 kpc with a minimum level of refinement $\ell_{\text{min}} = 8$ and a maximum level of refinement $\ell_{\text{max}} = 16$ ($\ell_{\text{max}} = 17$) for the low (high) resolution run, resulting in a maximum spatial resolution of 7 pc (resp. 3.5 pc).

The mesh in Eulerian codes is static by definition. The quasi-Lagrangian AMR leads to an adaptive spatial resolution down to a minimum cell size. In contrast, in both moving mesh and Lagrangian codes, the resolution elements follow the motions of the gas, leading to a smoothly (and, in principle, infinitely) adaptive spatial resolution with a constant or near-constant
mass resolution. Despite being pseudo-Lagrangian, moving mesh codes share many properties with Lagrangian codes. For convenience, we will refer to GADGET, GIZMO, and AREPO as “Lagrangian codes” hereafter.

2.2. Initial Conditions

The initial conditions are generated with the MAKEDISKGALAXY code developed in Springel et al. (2005), which consists of a rotating disk galaxy embedded in a dark matter halo. The halo has a virial radius \( R_{\text{vir}} = 45 \) kpc and a virial mass \( M_{\text{vir}} = 10^{10} M_\odot \), and it follows a Hernquist profile, which matches an NFW (Navarro et al. 1997) profile at small radii with the concentration parameter \( c = 15 \) and the spin parameter \( \lambda = 0.035 \). The baryonic mass fraction is 0.8\%, out of which 10\(^5\) \( M_\odot \) is in the stellar disk and 7 \times 10\(^7\) \( M_\odot \) in the gaseous disk. Both the stellar and gaseous disks follow an exponential density for numerical purposes, which is uniformly distributed in the background. In contrast, there is no background gas in the halo for GIZMO and GADGET.

2.3. Numerical Resolution

For comparison purposes, we should, ideally, adopt the same resolution in both Lagrangian and Eulerian codes. However, Lagrangian codes have a fixed mass resolution and an adaptive spatial resolution while Eulerian codes are the opposite. Due to this intrinsic difference, there is no unique way of choosing the same resolution for both methods. In this work, we adopt a spatial resolution of \( \Delta x = 7 \) pc in the low-resolution model for the Eulerian code RAMSES. At this resolution, the thermal Jeans length can be resolved up to \( n \sim 100 \) cm\(^{-3} \) assuming \( T = 30 \) K. Therefore, for the Lagrangian codes, we choose the gas particle mass \( m_g \) such that the smoothing length is comparable to \( \Delta x \) at \( n = 100 \) cm\(^{-3} \), which translates to \( m_g = 100 \) \( M_\odot \) in the low-resolution models. The gravitational softening for the baryons (i.e., gas, preexisting stars, and the stars formed in the simulation) is set to be \( h_g = 10 \) pc. The dark matter halo is resolved with a much coarser resolution with the particle mass \( m_{dm} = 10^4 \) \( M_\odot \) and the gravitational softening \( h_{dm} = 200 \) pc. In the high-resolution models, we simply decrease all the spatial resolutions by a factor of 2 and the mass resolutions by 8. Table 1 summarizes the resolution-related parameters.

2.4. Radiative Cooling

We use the public GRACKLE library (Smith et al. 2017)\(^\text{12} \) for radiative cooling, adopting its equilibrium cooling table without solving a chemistry network. The gas metallicity is \( Z = 0.002 (0.1 Z_\odot) \) and is constant throughout the simulation.

\(^\text{12} \) https://grackle.readthedocs.io/

---

**Table 1: Numerical Resolution**

| Resolution | \( m_{dm} \) (\( 10^4 \) \( M_\odot \)) | \( m_g \) (\( M_\odot \)) | \( h_{dm} \) (pc) | \( h_g \) (pc) | \( \Delta x \) (pc) |
|------------|------------------|-----------------|-----------------|-----------------|-----------------|
| low        | 10               | 100             | 200             | 10              | 7               |
| high       | 1.25             | 12.5            | 100             | 5               | 3.5             |

*Note. \( m_{dm} \): particle mass of dark matter. \( m_g \): particle mass of gas. \( h_{dm} \): gravitational softening length for dark matter. \( h_g \): gravitational softening length for gas. \( \Delta x \): cell size.*

In addition, we include heating from the photoelectric effect. Following Bakes & Tielens (1994) and Wolfire et al. (2003), the photoelectric heating rate can be expressed as

\[
\Gamma_{\text{PE}} = 1.3 \times 10^{-24} n \epsilon_{\text{PE}} G_0 Z_\odot^2 \text{erg cm}^{-3} \text{s}^{-1},
\]

where \( n \) is the hydrogen number density, \( \epsilon_{\text{PE}} \) is the photoelectric heating efficiency, \( G_0 \) is the radiation strength in units of the Habing field (Habing 1968), and \( Z_\odot \) is the dust-to-gas mass ratio relative to the Milky Way value (\( \sim 1\% \)). We adopt a constant \( \epsilon_{\text{PE}} \) of 0.05 (see Appendix A) and \( Z_\odot = 0.1 \), assuming a linear \( Z-Z_\odot \) relationship.\(^\text{13} \) We assume \( G_0 \) scales linearly with the SFR surface density (\( \Sigma_{\text{SFR}} \)), normalized to the solar-neighbourhood values of \( G_0 = 1.7 \) and \( \Sigma_{\text{SFR}} = 2.4 \times 10^{-3} M_\odot \text{yr}^{-1} \text{kpc}^{-2} \). Since the SFR is unknown prior to the simulation, we use low-resolution simulations to empirically determine the photoelectric heating rate as a function of the galactocentric radius \( R \):

\[
\Gamma_{\text{PE}} = \Gamma_{\text{PE,0}} \exp \left( \frac{R_c - \max(R, R_c)}{R_c} \right)
\]

where \( \Gamma_{\text{PE,0}} = 2.6 \times 10^{-27} \) erg s\(^{-1} \) and \( R_c = 0.4 \) kpc. Our low-resolution experiments suggest that the results are insensitive to modest variations of \( \Gamma_{\text{PE}} \) (a factor of 3), which is consistent with Hu et al. (2017) and Smith et al. (2021).

2.5. Star Formation

We adopt the stochastic star formation recipe commonly used in simulations of galaxy formation. The local SFR is \( \rho_{\text{SF}} = \epsilon_{\text{SF}} \rho_{\text{gas}} h_{\text{ff}} / t_{\text{ff}} \), where \( \rho_{\text{gas}} \) is the gas density, \( \epsilon_{\text{SF}} \) is the star formation efficiency, and \( t_{\text{ff}} = \left( 3\pi/32 G_0 \rho_{\text{gas}} \right) \) is the free-fall time, and \( G \) is the gravitational constant. We adopt \( \epsilon_{\text{SF}} = 1\% \) as our fiducial choice, but we also explore \( \epsilon_{\text{SF}} = 100\% \) in some models. Instead of adopting a density threshold for star formation, we allow star formation only to occur when the thermal Jeans length \( L_J = \sqrt{\pi c_s^2 (G \rho_{\text{gas}})} < L_{J,0} \), where \( c_s \) is the sound speed and \( L_{J,0} \) is a predefined star formation threshold. We adopt \( L_{J,0} = \Delta x \) such that gas is eligible for star formation only when the Jeans length becomes unresolved as we can no longer follow the gravitational collapse faithfully (Truelove et al. 1997). For Lagrangian codes, this star formation criterion roughly corresponds to the density where the Jeans mass becomes unresolved (Bate & Burkert 1997).

Star particles are stochastically created based on the local SFR. In the Lagrangian codes, the gas resolution element is

\(^\text{13} \) Observations suggest that the linear relationship between \( Z \) and \( Z_\odot \) breaks down at low metallicity (Rémy-Ruyer et al. 2014). However, its impact on the thermal balance in the ISM can be rather small if heating from SN feedback dominates over photoelectric heating (Hu et al. 2016, 2017).
2.6. Supernova Feedback

We include feedback from core-collapse SNe. For a typical stellar population, there is about one SN progenitor in every 100 $M_\odot$ stellar mass. As our star particle has a mass of $m_g$ exactly, while star particles in AREPO are guaranteed to have a mass within a factor of 2 of $m_g$. In RAMSES, a star particle with a mass of $m_g$ is spawned by removing a gas mass of $m_g$ from the parent cell.

2.7. Models

The models we run are summarized in Table 2. For the low-resolution case, we only run our fiducial model where $\epsilon_{\text{SF}} = 1\%$ and $t_{\text{SN}} = 10$ Myr. For the high-resolution case, we run our fiducial model, a model with $t_{\text{SN}} = 0$, and a model with $\epsilon_{\text{SF}} = 100\%$. For RAMSES, we run an additional model with $t_{\text{SN}} = 50$ Myr. Finally, we run a model for the Lagrangian codes (GIZMO, AREPO, and GADGET), where we adopt $t_{\text{SN}} = 0$ up to $t = 50$ Myr and $t_{\text{SN}} = 10$ Myr afterward. As we will demonstrate, this particular setup turns out to be our favored fiducial model for Lagrangian codes as it prevents the artificial blowout of the disk during the initial phase. We do not have the GADGET runs for all the models (i.e., there is no gadget and gadget_sfe100).

3. Results

3.1. Star Formation Rate

Figure 1 shows the star formation rate (SFR) as a function of time for the low- and high-resolution models in the top and bottom panels, respectively. The average SFRs in the low-resolution models are in broad agreement in all codes. However, the ramses_lr model is notably less bursty (i.e., the SFR varies less with time) than the Lagrangian models. In addition, there is an initial starburst in the Lagrangian models at $t \sim 100$ Myr that is absent in the ramses_lr model. In the high-resolution models (bottom panel), the SFRs in the Lagrangian models become even burstier. Indeed, the initial starburst leads to extremely energetic SN feedback, which blows out the entire gaseous disk. Although a quasi-steady state is eventually established after some of the blown-out gas falls back, the overall gas surface density becomes significantly lower than that in the initial conditions, leading to a reduced time-averaged SFR. The blowout, however, is a somewhat artificial consequence of our initial conditions and numerical setup. Before the first star formation occurs, gas cools and collapses globally into an artificially thin disk. Once the gas becomes dense enough to form stars, the resulting coherent SNe can easily break out of the thin disk, resulting in a catastrophic blowout. This is a well-known phenomenon in simulations of isolated galaxies as well as in ISM patches, and several methods have been adopted to mitigate the initial blowout, such as turbulent driving (e.g., Walch et al. 2015; Kim & Ostriker 2017) or random SN driving (e.g., Hu et al. 2017; Smith et al. 2021). On the other hand, the ramses model still shows a nonbursty SFR similar to its low-resolution counterpart and there is no initial blowout of the disk, in sharp contrast to the Lagrangian models. An important clue in understanding this difference comes from the observation that the SN delay time has a significant effect on the burstiness of star formation in the Lagrangian codes. The top panel of Figure 2 shows the SFR time evolution

| Model Name | Code      | Resolution | $\epsilon_{\text{SF}}$ | $t_{\text{SN}}$ (Myr) | Additional Comments |
|------------|-----------|------------|-------------------------|------------------------|---------------------|
| gizmo_lr   | GIZMO     | 100 $M_\odot$ | 1%                      | 10                     |                     |
| arepo_lr   | AREPO     | 100 $M_\odot$ | 1%                      | 10                     |                     |
| gadget_lr  | GADGET    | 100 $M_\odot$ | 1%                      | 10                     |                     |
| ramses_lr  | RAMSES    | 7 pc       | 1%                      | 10                     |                     |
| gizmo      | GIZMO     | 12.5 $M_\odot$ | 1%                      | 10                     |                     |
| arepo      | AREPO     | 12.5 $M_\odot$ | 1%                      | 10                     |                     |
| ramses     | RAMSES    | 3.5 pc     | 1%                      | 10                     |                     |
| gizmo_tsn0 | GIZMO     | 12.5 $M_\odot$ | 1%                      | 0                      |                     |
| arepo_tsn0 | AREPO     | 12.5 $M_\odot$ | 1%                      | 0                      |                     |
| ramses_tsn | RAMSES    | 3.5 pc     | 1%                      | 0                      |                     |
| ramses_tsn50 | RAMSES   | 3.5 pc     | 1%                      | 50                     |                     |
| gizmo$s^*$ | GIZMO     | 12.5 $M_\odot$ | 1%                      | 10                     | $t_{\text{SN}} = 0$ until $t = 0.25$ Gyr |
| arepo$s^*$ | AREPO     | 12.5 $M_\odot$ | 1%                      | 10                     | $t_{\text{SN}} = 0$ until $t = 0.25$ Gyr |
| gadget$s^*$ | GADGET   | 12.5 $M_\odot$ | 1%                      | 10                     | $t_{\text{SN}} = 0$ until $t = 0.25$ Gyr |
| gizmo_sfe100 | GIZMO    | 12.5 $M_\odot$ | 100%                    | 10                     |                     |
| arepo_sfe100 | AREPO   | 12.5 $M_\odot$ | 100%                    | 10                     |                     |
| ramses_sfe100 | RAMSES | 3.5 pc     | 100%                    | 10                     |                     |

Note. Models with an asterisk are run with $t_{\text{SN}} = 0$ for the first 0.25 Gyr to prevent the initial artificial blowout of the gaseous disk. $\epsilon_{\text{SF}}$: star formation efficiency. $t_{\text{SN}}$: SN delay time.
for models with instantaneous SN feedback ($t_{SN} = 0$). With this change, the SFRs in gizmo\_tsn0, ramses\_tsn0, and gadget\_tsn0 all show excellent agreement with each other. The SFR in arepo\_tsn0 is lower by a factor of 2, and it is unclear what causes the slight difference. More importantly, the SFRs in all codes are strikingly constant with little temporal fluctuation. In addition, the initial burst is absent and a quasi-steady state is rapidly established without the blowout of the disk.

An interesting implication is that instantaneous SN feedback can be a simple alternative method to mitigate the artificial initial starburst. Motivated by this, we rerun our Lagrangian codes with $t_{SN} = 0$ for $t < 250$ Myr and $t_{SN} = 10$ Myr for $t \geq 250$ Myr. We do not do so with RAMSES as it shows no initial starburst. The results are shown in the middle panel of Figure 2. The Lagrangian models show reasonable agreement with each other. Although there is still a burst of SFR at $t \geq 250$ Myr right after we switch $t_{SN}$ from 0 to 10 Myr, it is much weaker and does not blow out the disk. Consequently, the time-averaged SFR is not reduced and is comparable to ramses. We therefore refer to these models as our “fiducial” Lagrangian models. However, even if the artifacts of the initial conditions have been greatly reduced, the SFRs for the Lagrangian codes are still burstier than ramses. This suggests that the difference in burstiness between the two methods is an intrinsic property rather than an artifact of the initial conditions.

The observed SFR in the WLM galaxy from Hunter et al. (2010) is $1.7 \times 10^{-3} M_\odot$ yr$^{-1}$ from H$\alpha$ and $6.3 \times 10^{-3} M_\odot$ yr$^{-1}$ from far-ultraviolet (FUV). The SFRs in our fiducial models are broadly in good agreement with observations, which is reassuring.

The bottom panel of Figure 2 shows models with $\epsilon_{SF} = 100\%$ as well as a RAMSES model with $t_{SN} = 50$ Myr. Both gizmo\_sfe100 and arepo\_sfe100 show very bursty SFRs similar to the corresponding models gizmo and arepo where $\epsilon_{SF} = 1\%$. On the other hand, both ramses\_sfe100 and ramses\_tsn50 are bursty, in contrast to ramses. In other words, the burstiness in Eulerian codes can be enhanced by increasing either $\epsilon_{SF}$ or $t_{SN}$. We will explore the reason for this in later sections.

The time-averaged SFRs of different models are summarized in Table 3.
3.2. Gas Morphology and Thermal State

Figure 3 shows the face-on maps of the gas surface density in different models at $t = 500$ Myr. The most striking feature is that large SN-driven bubbles are clearly seen in some models but are completely absent in others. In fact, the presence of SN bubbles is closely correlated with the burstiness of the SFR shown in Figures 1 and 2: Models with burstier star formation show larger SN bubbles. In particular, those that experience an initial blowout have the largest SN bubbles, and their gas surface densities are notably reduced even at $t = 500$ Myr, as some of the blown-out gas never falls back to the disk. On the other hand, SN bubbles are completely absent in models where $\dot{M}_{\text{SN}} = 0$. Similarly, the standard $\text{ramses}$ runs show very few SN bubbles, consistent with its nonbursty star formation history. SN bubbles in RAMSES models can only be generated by increasing either $\epsilon_{\text{SR}}$ ($\text{ramses}_\text{sfe100}$) or $t_{\text{SN}}$ ($\text{ramses}_\text{tsn50}$).

Figure 4 shows the two-dimensional mass-weighted normalized distribution of the hydrogen number density versus temperature (the so-called "phase diagram") at $t = 500$ Myr. The red solid line indicates the star formation threshold where $L_{\text{j}} = 3.5$ pc. Broadly speaking, all models show similar distributions in the phase diagram despite their differences in star formation burstiness and gas morphology. This is because the thermal balance in the ISM is mainly controlled by radiative cooling. The majority of gas is in the diffuse warm gas, where $n \sim 0.1$ cm$^{-3}$ and $T \sim 10^4$ K, which is consistent with Hu et al. (2016, 2017). However, models with $t_{\text{SN}} = 0$ produce less hot ($T > 10^5$ K) and diffuse gas compared to their fiducial counterparts. This is expected as the hot and diffuse gas typically exists in the SN bubbles that are largely absent in these models. The RAMSES runs have more gas in the density range of $n \sim 10–100$ cm$^{-3}$ as indicated by the slightly brighter color but have very little gas above the star formation threshold.

3.3. Environments for Supernovae and Star Formation

The environment where SNe occur provides crucial information on the efficiency of SN feedback. The left panel of Figure 5 shows the cumulative distribution of the ambient gas density where SNe occur ($n_{\text{SN}}$) for the Lagrangian models. The agreement between different Lagrangian codes is remarkable. Most SNe occur in diffuse gas where $n_{\text{SN}} \sim 10^{-2}$ cm$^{-3}$ in the fiducial models ($\text{gizmo}^*$, $\text{arepo}^*$, and $\text{gadget}^*$) and therefore are well resolved. In contrast, most SNe occur in dense gas where $n_{\text{SN}} \sim 10^2$ cm$^{-3}$ in the instantaneous SN models ($\text{gizmo}_\text{tsn0}$, $\text{arepo}_\text{tsn0}$, and $\text{gadget}_\text{tsn0}$). This is expected, as SN feedback kicks in right after star formation, which by construction only occurs in cold and dense gas, leading to more efficient energy loss from radiative cooling, which explains the nonburstiness and the lack of SN bubbles. Although these SNe are formally unresolved, the time-averaged SFRs in these models are still comparable to those in the fiducial models, suggesting that these “unresolved” SNe are still able to inject enough momentum to regulate star formation.

In contrast, $\text{ramses}$ shows a very different $n_{\text{SN}}$ distribution compared to the Lagrangian codes with a significant fraction of SNe occurring at $n > 10$ cm$^{-3}$, as shown in the right panel of Figure 5. This is part of the explanation for why $\text{ramses}$ behaves more like the instantaneous SN models. However, there are still $\sim 25\%$ of SNe with $n_{\text{SN}} < 0.01$ cm$^{-3}$ in $\text{ramses}$, which is not significantly lower compared to $\text{gizmo}^*$. This implies that SNe occurring at low densities is not a sufficient condition for efficient feedback, which we will discuss in more detail in Section 3.4. On the other hand, both $\text{ramses}_\text{sfe100}$ and $\text{ramses}_\text{tsn50}$ have a large fraction of low-$n_{\text{SN}}$ SNe and therefore show buoyant SFRs and large SN bubbles.

Having established that the burstiness and gas morphology are both related to $n_{\text{SN}}$, we now need to understand what leads to the difference in $n_{\text{SN}}$ in different models.

The left panel of Figure 6 shows the cumulative distribution of the gas density where star formation occurs ($n_{\text{SF}}$) for the Lagrangian codes. The agreement between different Lagrangian codes is again remarkable. With instantaneous SN feedback (dashed lines), most star formation occurs in a narrow range of densities where $n \sim 500$ cm$^{-3}$. Once the gas reaches high enough densities to form stars, SN feedback that occurs instantaneously is able to stop the gas from further collapsing. Therefore, stars always form around the star

Note. (4) Time-averaged SFR in units of $10^{-3}$ $M_\odot$ yr$^{-1}$. (5) Time-averaged mass-loading factor. (6) Time-averaged energy-loading factor. The first 250 Myr is excluded from the time-averaging to discard the initial transient phase.
formation threshold density. In contrast, in the fiducial models with $t_{SN} = 10$ Myr (solid lines), star formation occurs over a large range of densities, spanning more than two orders of magnitude from $500$ to $10^3$ cm$^{-3}$. In this case, gas keeps collapsing toward higher densities even after it exceeds the threshold density, as there is no countering mechanism against gravity before the first SNe occur in 10 Myr. In contrast, as shown in the right panel of Figure 6, *ramses* behaves very differently from its Lagrangian counterparts, showing a narrow range of $n_{SF}$ very similar to the instantaneous SN models.

We therefore conclude that Lagrangian and Eulerian codes behave very differently at densities above the star formation threshold where the Jeans length becomes unresolved: While gas in Lagrangian codes continues to collapse to much higher (unresolved) densities, gas in Eulerian codes lingers around the star formation threshold. This is because the spatial resolution in Lagrangian codes is adaptive and therefore it is much easier for the unresolved gas to collapse. In comparison, the collapse of unresolved gas in Eulerian codes is limited by the fixed spatial resolution and is thus relatively more difficult. Note that in both cases, the collapse is unresolved (and additional physical processes such as pre-SN feedback may become important) and so neither behavior is necessarily correct.

*Figure 3.* Gas surface density in the face-on view in different models at $t = 500$ Myr. The SN-driven bubbles are clearly visible in most cases except in models with $t_{SN} = 0$ and in *ramses*. 
3.4. Supernova Clustering

In this section, we study the clustering properties of SNe and show that SN clustering has a fundamental impact on \( n_{\text{SN}} \), SFR burstiness, and gas morphology. The effect of SN clustering is two-fold. First, it reduces the ambient gas densities such that the SN bubble will retain the overpressurized hot gas for a longer time as the radiative cooling rate becomes lower than the energy injection rate of the temporally clustered SNe (Gatto et al. 2015; Kim \& Ostriker 2015; Walch et al. 2015; Girichidis et al. 2016; Gentry et al. 2017). Second, it generates coherent gas flows with less momentum cancellation. Therefore, it facilitates the formation of large SN bubbles that can break out of the gaseous disk (Kim et al. 2017; Orr et al. 2022).

The SN clusters are defined by the friends-of-friends method following Smith et al. (2021). This is possible as we record the location and time of each SN event in the simulations. In addition to a linking length of 10 pc, we also impose a linking time of 1 Myr to ensure that the SNe are indeed temporally clustered. The number of SNe in each SN cluster is defined as the clustering number \( N_{\text{cl}} \). We perform the analysis in the “lab” frame and do not correct for the effect of galactic rotation. This means that an SN might drift away from the SN cluster it physically belongs to by more than a linking length due to rotation and thus it would be incorrectly excluded. We argue that this should only be a minor effect as most of our SNe are temporally clustered at much shorter timescales than 1 Myr (see Figure 10).

Figure 7 shows the spatial distributions of the top 10 SN clusters (ranked by \( N_{\text{cl}} \)) in different models. Each point represents an SN event, and different colors represent different SN clusters. The average clustering number of the top 10 SN clusters \( \langle N_{\text{cl}} \rangle \) is shown in each panel.

Comparing the fiducial models, \textit{gizmo} and \textit{arepo} show significantly more clustered SNe than \textit{ramses}, with \( \langle N_{\text{cl}} \rangle \) larger by about an order of magnitude. In Lagrangian codes, as gas is able to collapse to densities far above the threshold density, star formation is locally enhanced due to the \( n_{\text{SN}} \) dependency. This makes the SNe highly clustered, which collectively generates large SN bubbles and low \( n_{\text{SN}} \), and leads to bursty star formation. The significantly less clustered SNe in \textit{ramses} suggests that having a low \( n_{\text{SN}} \), as we showed in Figure 5, is not a guarantee for efficient feedback—the SNe have to be clustered, too. Meanwhile, the location of SNe in \textit{gizmo} and \textit{arepo} closely follows the trajectory of galactic rotation, implying that the SN progenitors rarely drift away from their birth clouds and explode in the diffuse medium, which is a potential cause for the low \( n_{\text{SN}} \). This is perhaps not surprising as we do not resolve the collisional stellar dynamics and neither do we include any subgrid treatment for the so-called “runaway” or “walkaway” stars (de Mink et al. 2014).

To further support our argument that strong SN clustering is indeed caused by locally enhanced SFR, we conduct another numerical experiment, presented in Appendix B, where we force the local SFR to scale linearly rather than superlinearly with gas density. In this case, despite the formation of dense clouds, star formation in the clouds proceeds slowly, resulting in low clustering of SNe and an attendant lack of burstiness (resulting in inefficient outflows)—see the following section. In this way, the Lagrangian code behaves similarly to the fiducial Ramzes model despite using the fiducial SN delay time of 10 Myr, emphasizing that the root difference is tied ultimately to SN clustering.

On the other hand, the instantaneous SN models (\textit{gizmo}_{tsn0}, \textit{arepo}_{tsn0}, and \textit{ramses}_{tsn0}) represent the extreme case where SN clustering is strongly suppressed. Once star formation occurs, SN feedback immediately stops any further gravitational collapse and the development of any subsequent SNe. This leads to high \( n_{\text{SN}} \) and thus more rapid energy loss, as well as smaller SN-driven bubbles, resulting in a weaker dynamical impact on the ISM.

With \( \xi_{\text{SF}} = 100\% \), SNe are highly clustered in all models. The Lagrangian codes show a similar level of SN clustering compared to their fiducial counterparts (where \( \xi_{\text{SF}} = 1\% \)). The lack of clusters in the central area of \textit{arepo}_{sfe100} is caused by its strong initial blowout. As opposed to the \textit{ramses} model, \textit{ramses}_{sfe100} shows significantly more clustered SNe, with \( \langle N_{\text{cl}} \rangle \) larger by a factor of 30, consistent with its low \( n_{\text{SN}} \), bursty SFR, and large SN bubbles.

Figure 8 shows \( N_{\text{cl}} \) as a function of the median time of the SNe in a cluster \( \langle t_{\text{cl}} \rangle \) in the \textit{gizmo} model. Each circle represents an SN cluster. The effect of the SN delay time on clustering is...
clearly demonstrated: once $n_{3SN}$ is switched from 0 to 10 Myr at $t = 250$ Myr (indicated by the vertical dashed line), SNe immediately become significantly clustered.

To demonstrate the development of SN bubbles, Figure 9 shows the SN ambient temperature ($T_{SN}$, left panel) and density ($n_{SN}$) as a function of the normalized time ($t - t_{SN}$) for the top four SN clusters (ranked in $N_{cl}$) in the gizmo$^3$ model. Each point represents an SN event and different colors and symbols represent different clusters. All of these clusters share a similar time evolution: the first SNe occur in cold and dense gas, which gradually heat up and evacuate the gas, eventually saturating at $T_{SN} \sim 10^8$ K and $n_{SN} \sim 10^{-3}$ cm$^{-3}$.

To quantify the environmental properties of SN clusters, we calculate the median of the ambient temperature ($T_{cl}$) and density ($n_{cl}$) of each SN cluster as a function of $t_{cl}$ in Figure 10. The solid line shows the median values of $T_{cl}$ or $n_{cl}$ in each $N_{cl}$ bin while the shaded area brackets the 25th and 75th percentiles. For gizmo$^3$ and arepo$^3$, SNe are forced to occur in cold and dense gas by construction. In contrast, both gizmo and arepo, as $N_{cl}$ increases, $T_{cl}$ increases while $n_{cl}$ decreases. At $N_{cl} > 20$, $T_{cl}$ becomes higher than $10^5$ K, indicating the development of hot gas in the SN bubbles, which has been shown to be critical for launching galactic outflows (Walch et al. 2015; Girichidis et al. 2016; Hu 2019).

### 3.5. Galactic Outflows

Galactic outflows driven by SNe play a critical role in galaxy formation. We characterize the outflows in a similar fashion as in Hu (2019). We define the mass outflow rate as

$$M_{out} = \sum_i \frac{m_{g,i} v_{z,i}}{\Delta z},$$

and the energy outflow rate as

$$E_{out} = \sum_i \frac{m_{g,i} (v_i^2 + \gamma u_i) v_{z,i}}{\Delta z},$$

where $v_i$ is the gas velocity, $v_{z,i}$ is the gas velocity along the vertical direction, $\gamma = 5/3$ is the adiabatic index, $u_i$ is the specific internal energy, and $\Delta z = 0.1$ kpc is the thickness of the plane where outflows are measured. Here $z = 0$ corresponds to the midplane of the disk in the initial conditions. The summation is over cells or particles that are (i) located between $z = 10 \pm 0.5 \Delta z$ kpc or $z = -10 \pm 0.5 \Delta z$ kpc and (ii) traveling “outward,” i.e., $v_{z,i} > 0$.

We define the mass-loading factor,

$$\eta_m = \frac{M_{out}}{\langle SFR \rangle},$$

and the energy-loading factor,

$$\eta_e = \frac{E_{out}}{E_{SN} \langle SFR \rangle},$$

where $E_{SN} = 10^{51}$ erg is the injection energy per SN and $M_{SN} = 100 M_\odot$ is the corresponding mass formed in a stellar population per SN. The bracket ($\langle \rangle$) represents time-averaging, and we exclude $t < 250$ Myr to discard the initial transient phase. We use the time-averaged SFR instead of the instantaneous SFR as the normalization factor such that the fluctuations of $\eta_m$ and $\eta_e$ are purely due to the fluctuations in the outflows. The time-averaged $\langle \eta_m \rangle$ and $\langle \eta_e \rangle$ summarized in Table 3, are therefore

$$\langle \eta_m \rangle = \langle M_{out} \rangle / \langle SFR \rangle$$

and

$$\langle \eta_e \rangle = M_{SN} \langle E_{out} \rangle / E_{SN} \langle SFR \rangle,$$

respectively.

Figure 11 shows $M_{out}$ (top panel) and $E_{out}$ (bottom panel) as a function of time across the planes of $z = \pm 10$ kpc for the Lagrangian models. Instantaneous SN models have more than two orders of magnitude lower outflow rates compared to the fiducial models, demonstrating that SN clustering substantially enhances outflows. Both gizmo$^3$ and arepo$^3$ show weak but nonzero outflows at $t \sim 100$ Myr as a result of the initial star formation, which is absent in arepo$^3$. This is
likely due to the difference in the initial conditions: arepo_tsn0 includes a low-density background for numerical purposes while gadget_tsn0 and gizmo_tsn0 adopt a vacuum background that presumably facilitates the relatively weak outflows to expand and fill up the vacuum.

Interestingly, the outflow rates in the fiducial models show notable differences even though their SFRs, \eta_{SN}, and \eta_{SF} are all in good agreement. The difference between arepo* and gadget* is mostly due to their SFRs, which differ by a factor of 3, while their \langle \eta_{hl} \rangle and \langle \eta_{k} \rangle only differ by a factor of 25% and 80%, respectively. However, gizmo* is a factor of 4–5 lower in \langle \eta_{hl} \rangle and a factor of 10–18 lower in \langle \eta_{k} \rangle compared to arepo* and gadget*. The origin of this difference is still unclear and requires further investigation in future work.

We now compare outflows between Eulerian and Lagrangian codes in Figure 12. For clarity, we only show the AREPO code. The GADGET code. Gutcke et al. (2021) found notable differences even though their SFRs, \eta_{SN}, and \eta_{SF} are all in good agreement. The difference between arepo* and gadget* is mostly due to their SFRs, which differ by a factor of 3, while their \langle \eta_{hl} \rangle and \langle \eta_{k} \rangle only differ by a factor of 25% and 80%, respectively. However, gizmo* is a factor of 4–5 lower in \langle \eta_{hl} \rangle and a factor of 10–18 lower in \langle \eta_{k} \rangle compared to arepo* and gadget*. The origin of this difference is still unclear and requires further investigation in future work.

We conclude that SN clustering plays a fundamental role in driving galactic outflows. Only models with significant SN clustering show \langle \eta_{hl} \rangle larger than unity.

4. Discussion

4.1. Comparison with Previous Works

Several recent studies have quantified outflows in terms of \eta_{hl} and \eta_{k} in similar initial conditions of a dwarf galaxy with resolved SN feedback. Our fiducial Lagrangian models (gizmo*, arepo*, and gadget*) are in broad agreement with these studies: Hu (2019) found \langle \eta_{hl} \rangle \sim 4 and \langle \eta_{k} \rangle \sim 0.07 at \mid z \mid = 10 \, kpc using the GADGET code. Smith et al. (2021) found \eta_{hl} fluctuating between 1 and 10 and \eta_{k} between 0.003 and 0.3 at \mid z \mid = 10 \, kpc in a comparable model (their FIXED_SN_ENERGY) using the AREPO code. Gutcke et al. (2021) found \eta_{hl} between 5 and 10 and \eta_{k} between 10^{-4} and 10^{-3} at \mid z \mid = 2 \, kpc in a comparable model (their FIXED_SN_ENERGY) using the AREPO code. We note that \eta_{hl} is in better agreement among these studies than \eta_{k}, probably because \eta_{k} is very sensitive to small differences in the gas velocity (see Equation (4)). Our results confirm these previous findings that the predicted \eta_{hl} from SN-resolved galaxy-scale simulations are significantly lower than what cosmological simulations commonly assume in their subgrid models. Observations of nearby dwarf galaxies seem to support the low-\eta_{hl} case (McQuinn et al. 2019).

In a similar setup, Hu et al. (2017) found that a shortened SN delay time (\tau_{SN} = 3 \, Myr) has the effect of suppressing the formation of SN bubbles and galactic outflows, which is consistent with our results. However, they did not find an effect of \tau_{SN} on the burstiness of star formation. Smith et al. (2021) showed that photoionization from massive stars prior to the SN events (the so-called “early feedback”) reduces SN clustering and thus significantly suppresses the burstiness of star formation, formation of SN bubbles, and galactic outflows, reducing \eta_{hl} and \eta_{k} by more than an order of magnitude. This is very similar to our instantaneous SN models, which can be viewed as an extreme case of early feedback. Keller & Kruisjes | 2022 simulated a Milky Way–like galaxy with \msun \sim 10^{12} \, M_{\odot}, where SN feedback is modeled in a subgrid fashion. Although the simulated galaxy and the numerical treatments are very different, they found that a longer SN delay time enhances galactic outflows by enhancing the clustering of young stars, which is qualitatively consistent with our results.

Kim et al. (2016) have conducted a detailed code comparison study for an isolated Milky Way–like galaxy with significantly more participating codes. They found that different codes generally agree well with each other. In particular, they did not find systematic differences between Lagrangian and Eulerian codes in terms of burstiness and the sizes of SN bubbles. This is probably because they have adopted the simple thermal injection as their subgrid prescription for SN feedback, which is known to be very inefficient at this resolution as most energy would be radiated away without generating much momentum.
4.2. Differences between Numerical Methods

As we have shown in Section 3.3, the differences between Lagrangian and Eulerian codes arise near the resolution limit when the Jeans length becomes unresolved. This is perhaps not surprising: Code differences almost by definition have to occur near the resolution limit, as the resolved scales should converge to the physical solutions for any numerically consistent method. However, in our case, the differences at the resolution limit quickly propagate to much larger, well-resolved scales due to clustered star formation and SN feedback.

As we alluded to in Section 3.3, once the gas enters the Jeans length-unresolved regime, both methods can no longer faithfully follow the collapse, and the evolution of gas depends sensitively on numerics such as gravitational softening. That said, it is interesting to ask which method is “less wrong.” For a collapsing cloud that becomes unresolved, both the Jeans mass and Jeans length would keep decreasing as the density increases. In other words, the cloud should keep collapsing and fragmenting. Lagrangian codes cannot capture the fragmentation without particle splitting or cell refinement, and they might underestimate the collapsing speed if the adopted softening length is much larger than the cell size. However, the cloud will continue to collapse as expected. In contrast, the collapse in an Eulerian code would halt when the Jeans length becomes unresolved as it requires accreting.
gas into the minimum cell. In this sense, Lagrangian codes are arguably more accurate than Eulerian codes in the unresolved regime.\(^{15}\)

The situation becomes more subtle if we consider the physical processes we do not include in this work. In particular, as we do not include feedback from photoionization, we might have overestimated SN clustering and more significantly so in our Lagrangian models. If this is the case, our Eulerian model would more accurately, though coincidentally, predict the SN clustering.

4.3. Star Formation Efficiency

Previous studies have shown that the star formation efficiency only has a weak effect on the galaxy-scale SFR in both Lagrangian and Eulerian codes (e.g., Agertz et al. 2013; Benincasa et al. 2016; Semenov et al. 2017; Hopkins et al. 2018). While this is consistent with our Lagrangian models, it appears to be in conflict with our Eulerian models where increasing $\epsilon_{\text{SF}}$ leads to a qualitative change in the burstiness of star formation due to enhanced SN clustering. This is probably because SN feedback is unresolved and treated in a subgrid fashion in those studies, which reduces its dynamical impact. Alternatively, as those studies simulated more massive galaxies, the local burstiness might have been averaged out such that the global SFR remains the same. In this case, we expect the outflow rates to increase with $\epsilon_{\text{SF}}$.

Instead of assuming a constant $\epsilon_{\text{SF}}$, some recent simulations of galaxy formation have adopted a class of subgrid models for star formation that calculate $\epsilon_{\text{SF}}$ based on local gas properties (Semenov et al. 2016; Kretschmer & Teyssier 2020). These models assume a subgrid log-normal density distribution as predicted by simulations of supersonic turbulence, and models differ in their criteria for the onset of star formation (Krumholz & McKee 2005; Hennebelle & Chabrier 2008; Padoan & Nordlund 2011; Federrath & Klessen 2012; Burkhart 2018; Calura et al. 2019; Calura et al. 2022).

\(^{15}\)Eulerian codes might behave more similarly to Lagrangian codes when adopting a minimum cell size well below the resolvable Jeans length. However, this would be prohibitively expensive in practical applications.

4.4. Numerical Convergence

The significant differences between our low- and high-resolution models suggest that we have not yet reached numerical convergence. This is consistent with Smith et al. (2018) whose models with $m_g = 20 M_o$ and $m_g = 200 M_o$ still show significantly different outflow rates. In addition, the convergence study in Hu (2019) with $m_g = 1$, $5$, $25$, and $125 M_o$ showed that convergence was only achieved at $m_g = 5 M_o$ when individual SNe are properly resolved. Therefore, we optimistically expect that our high-resolution models are actually close to convergence.

Many recent cosmological “zoom-in” simulations have adopted an SN feedback model that injects thermal energy for resolved SNe and injects the terminal momentum for unresolved SNe as a subgrid model (e.g., Hopkins et al. 2014, 2018; Agertz & Kravtsov 2015; Kimm et al. 2015; Marinacci et al. 2019). Most of these simulations have resolutions much coarser than our low-resolution models, suggesting that their SN feedback still operates in the subgrid regime where galactic outflows are expected to be underestimated due to the lack of hot gas (Hu 2019). More recent studies such as Agertz et al. (2020), Gutcke et al. (2022), and Calura et al. (2022) have started to resolve individual SNe in cosmological “zoom-in” simulations of dwarf galaxies, which is a promising way forward.

4.5. Missing Physics and Future Prospects

The main caveat of our results is the fact that we do not include pre-SN feedback such as photoionization. As we already discussed, this could lead to overestimated SN clustering, in particular for our Lagrangian models. This is a potential solution for the discrepancies we find between the two types of methods. An interesting follow-up project is therefore to include photoionization and see if the discrepancies can be mitigated. On the other hand, the suppression of SN clustering by photoionization in Smith et al. (2021) could also have been overestimated by the adopted “Strömgren-type” approach where dense clumps are preferentially ionized. Interestingly, Rathjen et al. (2021) conducted ISM-patch simulations using radiative transfer based on an inverse ray-tracing technique and reached the same conclusion that photoionization suppresses SN clustering, which leads to less efficient hot gas generation. This should be investigated by future galaxy-scale simulations using more accurate methods of radiative transfer such as adaptive ray tracing (Emerick et al. 2019). We note that including a solver for radiative transfer does not necessarily resolve the issue, as the important part is to ensure sufficient angular resolution to prevent the dense clumps from being preferentially ionized. This therefore implies a very demanding computational cost.

Another potentially important element is the collisional stellar dynamics. While we adopt a collisionless gravity solver in this work, the stellar dynamics in young star clusters is actually collisional. As our results highlight the importance of SN clustering, it is desirable to include an accurate $N$-body integrator, such as in Wall et al. (2020) and Rantala et al. (2021), to properly follows the evolution of young star clusters. However, the computational feasibility of incorporating it into galaxy-scale simulations remains to be explored. On the other hand, stellar
dynamics can be modeled in a subgrid fashion. Steinwandel et al. (2022) simulated a dwarf galaxy similar to ours, including a subgrid model for runaway massive stars that can enhance $\langle \eta_m \rangle$ by 50% and $\langle \eta_e \rangle$ by a factor of 5.

5. Summary

We have conducted a suite of simulations of an isolated dwarf galaxy using four different hydrodynamical codes (GIZMO, AREPO, GADGET, and RAMSES). All codes adopt the same physical model, which includes radiative cooling, photoelectric heating, star formation, and SN feedback. We directly resolve individual SN feedback without using subgrid models, which is a major source of uncertainty in cosmological simulations. Our main results can be summarized as follows.

1. The time-averaged SFRs and the distributions of gas density and temperature are in reasonable agreement in all codes (Table 3 and Figure 4). However, Lagrangian codes show a burstier star formation history (Figures 1 and 2), larger SN-driven bubbles (Figure 3), and stronger galactic outflows (Figures 11 and 12), in striking contrast to the Eulerian code. This originates from the different behaviors as gas collapses beyond the star formation threshold: the Jeans-length-unresolved gas collapses to much higher densities in the Lagrangian codes (Figure 6), leading to a more complete conversion of gas into stars and hence more highly clustered SNe (Figure 7). Hot gas ($T > 10^5$ K) is generated when $N_{cl} \gtrsim 20$. 

Figure 9. SN ambient temperature ($T_{SN}$, left) and density ($n_{SN}$, right) as a function of the normalized time ($t - t_{cl}$) for the top four SN clusters (ranked by $N_{cl}$) in the gizmo* model. Each point represents an SN event and different colors and symbols represent different clusters. The horizontal red line indicates the density above which the SN cooling radius becomes unresolved. The first SNe occur in cold and dense gas, which heat up and evacuate the gas, and the subsequent SNe occur in increasingly hot and diffuse gas.

Figure 10. Left: median SN ambient temperature of a cluster ($T_{cl}$) as a function of $N_{cl}$. The solid line shows the median values in each $N_{cl}$ bin while the shaded area brackets the 25th and 75th percentiles. Right: same as the left but for the median SN ambient density $n_{cl}$. The horizontal red line indicates the density above which the SN cooling radius becomes unresolved. Hot gas ($T > 10^5$ K) is generated when $N_{cl} \gtrsim 20$. 

The Astrophysical Journal, 950:132 (17pp), 2023 June 20 Hu et al.
2. If we let SN feedback occur with a zero delay time immediately after star formation as a numerical experiment, SN clustering would be strongly suppressed and SNe are forced to occur at high densities with rapid radiative losses. In this case, all codes behave similarly, showing a nonfluctuating SFR, no visible SN bubbles, and vanishing galactic outflows.

3. The adopted star formation efficiency ($\epsilon_{\text{SF}}$) has a significant effect on SN clustering in Eulerian codes, which in turn affects the star formation burstiness, sizes of SN bubbles, and outflow rates. In contrast, $\epsilon_{\text{SF}}$ only plays a minor role in Lagrangian codes where gas collapses to much higher densities such that local star formation is significantly enhanced, effectively enhancing $\epsilon_{\text{SF}}$ even when a low value is used.

4. Lagrangian models are in good agreement with each other in terms of gas morphology, SN densities, and star formation densities. However, GIZMO shows notably weaker outflows compared to AREPO and GADGET in the fiducial models, which requires further investigations in future work.

Acknowledgments

We thank the anonymous referee for the constructive comments that helped improve the manuscript, in particular for suggesting the additional run in Appendix B. We thank Volker Springel and Chang-Goo Kim for a stimulating discussion. C.Y.H. acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) via German-Israel Project Cooperation grant STE1869/2-1 GE625/17-1 and NASA ATP grant 80NSSC22K0716. The work of MCS was supported by a grant from the Simons Foundation (CCA 668771, LEH) and by the DFG under Germany’s Excellence Strategy EXC 2181/1-390900948 (the Heidelberg STRUCTURES Excellence Cluster). The Center for Computational Astrophysics is supported by the Simons Foundation. G.L.B. acknowledges support from the NSF (AST-2108470, XSEDE), a NASA TCAN award, and the Simons Foundation through their support of the Learning the Universe Collaboration. B.B. thanks the Alfred P. Sloan Foundation and the Packard Foundation for support as well as the support of NASA grant No. 80NSSC20K0500. Y.L. acknowledges financial support from NSF grant AST-2107735 and NASA grant 80NSSC22K0668.

Appendix A

Photoelectric Heating Efficiency

In this section, we investigate if our adopted constant photoelectric efficiency $\epsilon_{\text{PE}} = 0.05$ is a good approximation (see Equation (1)). In the upper-left panel of Figure 13, we show the gas temperature at thermal equilibrium as a function of density by running our adopted GRACKLE module for 1 Gyr at each density bin (green dotted line). The initial temperature is set at $10^4$ K and the metallicity is $0.1 Z_{\odot}$. In comparison, we also show the equilibrium temperature as a function of density obtained by running the nonequilibrium chemistry code developed by Simon Glover (Glover & Mac Low 2007; Glover & Clark 2012) (hereafter SGCHEM) for 1 Gyr at each density bin: The solid blue line is with a cosmic-ray (CR) ionization rate $\zeta = 10^{-16} \text{s}^{-1}$ motivated by $H_3^+$ observations in the Milky Way (Indriolo & McCall 2012) while the dashed orange line is with $\zeta = 0$.

The photoelectric efficiency adopted in SGCHEM is

$$\epsilon_{\text{PE}} = \frac{0.049}{1 + (0.004 \psi^{0.7})} + \frac{0.037 (T/10000)^{0.7}}{1 + 2 \times 10^{-4} \psi},$$

(A1)
where \( \psi = G_0 T^{0.5}/n_e \) and \( n_e \) is the electron number density, following Bakes & Tielens (1994), Wolfire et al. (2003), and Bergin et al. (2004). Here we have assumed that dust shielding is negligible, which is a fair approximation at low metallicity.

We find excellent agreement between GRACKLE and SGCHEM with \( \zeta = 10^{-16} \text{ s}^{-1} \) up to \( n \lesssim 10^2 \text{ cm}^{-3} \), which is reassuring. In contrast, SGCHEM with \( \zeta = 0 \) leads to a significantly lower equilibrium temperature for a broad range of densities. This is caused by a severely underestimated electron abundance \( x_e = n_e/n \) when CR ionization is switched off, as shown in the upper-right panel of Figure 13 (shown in orange), which in turn strongly underestimates \( \epsilon_{\text{PE}} \) (shown in blue).

The individual cooling and heating processes in SGCHEM are shown in the lower-left (for \( \zeta = 10^{-16} \text{ s}^{-1} \)) and lower-right (for \( \zeta = 0 \)) panels. For \( \zeta = 10^{-16} \text{ s}^{-1} \), heating is dominated by the photoelectric effect at \( n > 0.3 \text{ cm}^{-3} \) and by CR ionization at \( n < 0.3 \text{ cm}^{-3} \). The total heating rate is almost constant with \( n \), which is balanced by Ly\(\alpha \) cooling at low \( n \) and by fine-structure line cooling at high \( n \). At \( n > 10^3 \text{ cm}^{-3} \), heating from \( \text{H}_2 \) photodissociation and \( \text{H}_2 \) formation becomes non-negligible, which might explain the discrepancy in the equilibrium temperature in this regime. In contrast, for \( \zeta = 0 \), not only does CR ionization heating vanish by construction, but the photoelectric heating is also significantly suppressed, caused by the artificially low electron abundance. As a result, the equilibrium temperature becomes unreasonably low (\( T < 10^2 \text{ K} \)) even for typical densities in the diffuse ISM. This has been seen in Hu et al. (2017) where the authors adopted SGCHEM but did not include CR ionization. Interestingly, when applied to the hydrodynamical simulations of Hu et al. (2017), such low temperatures are never reached as SN feedback (and the turbulence it drives) provides extra heating and ionization, which has a similar effect to CR ionization. Therefore, the resulting phase diagram is quite similar to what we found in Figure 4. That said, having a more realistic thermal balance in a static medium is still desirable. We recommend future simulations that use SGCHEM (or other similar chemistry codes) where \( \epsilon_{\text{PE}} \) is calculated via Equation (A1) to always include CR ionization.

To conclude, adopting a constant \( \epsilon_{\text{PE}} \) is a good approximation of the more realistic situation where both the photoelectric effect and CR ionization are present, even though we do not explicitly include the latter. In reality, CR ionization either dominates heating at low densities or provides the crucial free electrons that facilitate photoelectric heating at high densities.

**Appendix B**

**Constant Star Formation Timescale**

In this section, we present a numerical experiment suggested by the anonymous referee that will further strengthen our conclusion on the importance of SN clustering. As we discussed in Section 3.4, the strong SN clustering in Lagrangian codes is due to the locally enhanced SFR from the superlinear density dependence \( \dot{\rho}_* = \epsilon_{\text{SF}} \rho_{\text{gas}}/t_{\text{SF}} \propto \rho_{\text{gas}}^{1.5} \). Here we conduct a simulation using GIZMO with \( f_{\text{SN}} = 10 \text{ Myr} \) but with the local SFR \( \dot{\rho}_* = \rho_{\text{gas}}/t_{\text{SFR}} \), where \( t_{\text{SFR}} = 2 \text{ Gyr} \) is...
the local depletion time, which is chosen to be a constant. In other words, the local SFR has a linear (rather than superlinear) dependence on gas density. As shown in Figure 14, this constant depletion time model shows a smooth gas morphology similar to the gizmo_tsn0 model but with more pronounced dense gas clumps. This is in contrast to the gizmo* model with large SN bubbles. In addition, as shown in Figure 15, this model shows a smooth, nonbursty SFR as a function of time and cumulative distributions of $n_{SF}$ and $n_{SN}$. The constant depletion time model behaves more similarly to the gizmo_tsn0 model—both have nonbursty star formation and vanishing outflows as a result of low SN clustering.

threshold density, the local SFR is not enhanced, leading to a large number of dense clumps. These clumps, instead of quickly converting into stars as in the gizmo* model, simply wander around, forming stars at a much lower rate, until they get dispersed by SN feedback after 10 Myr. This leads to low SN clustering, with the maximum $N_{cl} = 5$ in the entire simulation, which in turn leads to vanishing outflows.

These results support our argument that the superlinear dependence of the local SFR on density $\rho_{k} \propto \rho_{gas}^{1.5}$ leads to strong SN clustering in Lagrangian codes. Adopting a constant local depletion time can make Lagrangian codes behave more like the
