

Entropy-Conserving Scheme for Modeling Nonthermal Energies in Fluid Dynamics Simulations

VADIM A. SEMENOV ^{1,*} ANDREY V. KRAVTSOV ^{2,3,4} AND BENEDIKT DIEMER ⁵

¹*Center for Astrophysics | Harvard & Smithsonian, 60 Garden St, Cambridge, MA 02138, USA*

²*Department of Astronomy & Astrophysics, The University of Chicago, Chicago, IL 60637, USA*

³*Kavli Institute for Cosmological Physics, The University of Chicago, Chicago, IL 60637, USA*

⁴*Enrico Fermi Institute, The University of Chicago, Chicago, IL 60637, USA*

⁵*Department of Astronomy, University of Maryland, College Park, MD 20742, USA*

ABSTRACT

We compare the performance of energy-based and entropy-conservative schemes for modeling non-thermal energy components, such as unresolved turbulence and cosmic rays, using idealized fluid dynamics tests and isolated galaxy simulations. While both methods are aimed to model advection and adiabatic compression or expansion of different energy components, the energy-based scheme numerically solves the non-conservative equation for the energy density evolution, while the entropy-conserving scheme uses a conservative equation for modified entropy. Using the standard shock tube and Zel'dovich pancake tests, we show that the energy-based scheme results in a spurious generation of nonthermal energy on shocks, while the entropy-conserving method evolves the energy adiabatically to machine precision. We also show that, in simulations of an isolated L_* galaxy, switching between the schemes results in ≈ 20 – 30% changes of the total star formation rate and a significant difference in morphology, particularly near the galaxy center. We also outline and test a simple method that can be used in conjunction with the entropy-conserving scheme to model the injection of nonthermal energies on shocks. Finally, we discuss how the entropy-conserving scheme can be used to capture the kinetic energy dissipated by numerical viscosity into the subgrid turbulent energy *implicitly*, without explicit source terms that require calibration and can be rather uncertain. Our results indicate that the entropy-conserving scheme is the preferred choice for modeling nonthermal energy components, a conclusion that is equally relevant for Eulerian and moving-mesh fluid dynamics codes.

Keywords: Hydrodynamical simulations; Galaxy evolution; Stellar feedback; Interstellar dynamics; Cosmic rays; Shocks

1. INTRODUCTION

Stellar feedback plays a critical role in shaping properties of simulated galaxies on all relevant scales, from the overall stellar mass and star formation rate to metallicity, galaxy morphology, and chemical abundances in the circumgalactic (CGM) and intergalactic medium (IGM) (e.g., Governato et al. 2010; Brook et al. 2012; Stinson et al. 2013; Hopkins et al. 2013, 2014; Agertz et al. 2013; Agertz & Kravtsov 2015, 2016, see also Naab & Ostriker 2017 and Vogelsberger et al. 2020 for recent reviews). Modeling feedback processes in numerical simulations of galaxy formation has been the focus of extensive theoretical efforts.

Over the past decades a number of different methods were developed to model stellar feedback, ranging from the direct injection of energy and momentum to the interstellar medium (ISM) when the physical scales of this injection can be resolved (e.g., Hopkins et al. 2011, 2012, 2017; Agertz et al. 2013; Marinacci et al. 2019; Smith et al. 2020; Jeffreson et al. 2021), to effective models of the ISM and outflow driving when the multi-phase ISM structure cannot be sufficiently resolved (e.g., Yepes et al. 1997; Springel & Hernquist 2003; Braun & Schmidt 2012; Vogelsberger et al. 2013; Springel et al. 2018).

One important class of feedback implementations is based on explicit modeling of a nonthermal energy component sourced by young stars in addition to gas thermal energy (e.g., Springel 2000; Teyssier et al. 2013; Agertz et al. 2013; Benincasa et al. 2016). Common examples of such nonthermal components are cosmic rays (e.g.,

* vadim.semenov@cfa.harvard.edu
NHFP Hubble Fellow.

Jubelgas et al. 2008; Uhlig et al. 2012; Booth et al. 2013; Salem & Bryan 2014; Salem et al. 2014; Pakmor et al. 2016; Ruszkowski et al. 2017; Wiener et al. 2017; Chan et al. 2019; Buck et al. 2020) and small-scale ISM turbulence (e.g., Braun et al. 2014; Schmidt et al. 2014; Semenov et al. 2016, 2021; Kretschmer & Teyssier 2020; Kretschmer et al. 2021). Such models differ by their sink, source, and extra transport terms (such as cosmic ray transport or turbulent diffusion). Without these terms, the behavior of the nonthermal energy is analogous to the thermal energy and consists of advection with gas density and the “ PdV ” work done during gas compression or expansion.

Implementation-wise, nonthermal energies can be modeled using the methods developed for modeling the thermal energy of the gas. Such methods were introduced to follow gas temperature in highly supersonic flows typical for galaxy formation simulations¹ where the modeling of thermal energy requires special attention. Indeed, when the flow is subsonic or only moderately supersonic, the thermal energy can be derived as the difference between the total and kinetic energy, $e_{\text{th}} = e_{\text{tot}} - e_{\text{kin}}$, which are readily available in conservative methods that follow the gas mass, momentum, and total energy. However, when the flow is highly supersonic, the difference $e_{\text{tot}} - e_{\text{kin}}$ becomes small and comparable to the truncation errors of e_{tot} and e_{kin} , requiring a more accurate method for modeling e_{th} .

Two such methods for modeling thermal energy have been introduced in the early days of hydrodynamic galaxy formation simulations. In the “dual-energy” formalism proposed by Bryan et al. (1995), thermal energy is modeled as a separate variable in addition to total energy, with the PdV source term being computed explicitly. The second method, proposed by Ryu et al. (1993), models the gas *entropy* as a separate variable, which then can be used to compute temperature or thermal energy. The key advantage of the latter method is that gas entropy obeys a conservative equation, at least outside of shocks and in regions of the flow where dissipative processes are negligible. The entropy equation can thus be solved accurately without explicit source terms that can introduce significant errors.

Both these methods involve several parameters including the criteria that are used to decide whether e_{th}

should be taken from the explicitly modeled variable or synchronized with the total energy as $e_{\text{th}} = e_{\text{tot}} - e_{\text{kin}}$ (Bryan et al. 1995; Ryu et al. 1993; Springel 2010; Teyssier 2015). In most idealized problems, both methods and different synchronization criteria work comparably well, and their differences tend to be small even in more realistic problems (e.g., Costa et al. 2020). However, in some cases, these choices can have significant effects. For example, Villasenor et al. (2020) showed that the formation of shock-heated gaseous halos in the IGM can be strongly delayed by the choice of the criteria used to synchronize thermal and total energies.

The choice of the method for modeling of *nonthermal* components may have an even more significant impact. Indeed, the special treatment of thermal energy is required only in the extreme regime of highly supersonic flows, which are typically limited to a small fraction of the problem volume and where thermal pressure is by definition negligible. In contrast, nonthermal energies must be followed separately from the total energy in the entire simulation domain. Therefore, the choice of the method can impact results via the dynamical effect of nonthermal pressure and via its effect on other physical processes—such as star formation, radiative cooling, chemistry, etc.—which are sometimes coupled to the nonthermal energy.

In this paper, we explore the impact of different choices in modeling nonthermal energy on the evolution of gas using a series of idealized tests and more realistic simulations of an isolated L_* galaxy. The paper is organized as follows. In Section 2, we review the methods of Bryan et al. (1995) and Ryu et al. (1993) and provide additional details about our implementation of these methods in the ART galaxy formation code. We also outline a simple method that can be used to implement the injection of nonthermal energy on shocks (Section 2.3). In Section 3, we compare the two methods in shock tube and hydrodynamic Zel’dovich pancake tests and show that the energy-based method results in the spurious generation of nonthermal energy on shocks while the entropy-conserving method can enforce the intended adiabatic behavior. Then, in Section 4, we further compare the methods in simulations of an isolated L_* galaxy and show that the choice of the method can significantly impact galaxy properties. In Section 5, we discuss our results and propose a novel method for modeling unresolved turbulent energy by capturing numerical dissipation with the entropy-conserving scheme. We summarize our conclusions in Section 6.

¹ For example, the motion of the cold ISM due to galactic rotation and translational velocity ($c_s \sim 1\text{--}3\text{ km s}^{-1}$, $v \sim 100\text{--}300\text{ km s}^{-1}$), accretion of cold gaseous streams at high redshifts, and other motions of $T \sim 10^4\text{ K}$ gas in the IGM ($c_s \sim 10\text{ km s}^{-1}$, $v \sim \text{few } 100\text{ km s}^{-1}$) are flows with extremely high Mach numbers, $v/c_s > 30$.

2. IMPLEMENTATION OF THERMAL AND NONTHERMAL ENERGIES

To explore the impact of specific choices made in modeling nonthermal energy components, we use idealized tests and more realistic simulations of an isolated galaxy performed with the adaptive mesh refinement (AMR) hydrodynamics and N -body code ART (Kravtsov 1999; Kravtsov et al. 2002; Rudd et al. 2008; Gnedin & Kravtsov 2011). The hydrodynamic fluxes in the ART code are handled by a second-order Godunov-type method (Colella & Glaz 1985) with a piecewise linear reconstruction of states at the cell interfaces (van Leer 1979) and a monotonized central slope limiter based on Colella (1985).

Here we summarize our implementation of different methods for solving the set of advection equations with PdV -work source terms for the thermal and nonthermal energy components. The key assumption is that the adiabatic evolution of these components can be modeled separately from any non-adiabatic processes, which can be added via source terms. Thus, during a hydrodynamics step and before the source terms are applied, the internal energies of the thermal and nonthermal components evolve adiabatically:

$$\frac{\partial e_{\text{th}}}{\partial t} + \nabla \cdot (e_{\text{th}} \mathbf{u}) = -P_{\text{th}} \nabla \cdot \mathbf{u}, \quad (1)$$

$$\frac{\partial e_{\text{nt}}}{\partial t} + \nabla \cdot (e_{\text{nt}} \mathbf{u}) = -P_{\text{nt}} \nabla \cdot \mathbf{u}. \quad (2)$$

Here and throughout the paper, e_{th} and e_{nt} denote energies per unit volume, and $\nabla \equiv (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ is the gradient operator. For clarity, we consider a single nonthermal energy component, e_{nt} , but all of the results and conclusions can be trivially generalized for an arbitrary number of such energy components.² In our calculations we assume an ideal gas equation of state for each of the components: $P_i = (\gamma_i - 1)e_i$ with $i \in [\text{th}, \text{nt}]$.

The above set of equations is coupled with gas dynamics via the total pressure $P = P_{\text{th}} + P_{\text{nt}}$. In addition, as thermal and nonthermal energies contribute to the gas total energy, e_{tot} , a suitable method should be adopted to ensure that $e_{\text{tot}} = e_{\text{kin}} + e_{\text{th}} + e_{\text{nt}}$, where $e_{\text{kin}} = \rho \mathbf{u}^2/2$, when such synchronisation is appropriate.

The choice of the synchronization method is dictated by the expected behavior of nonthermal energies across shocks. Indeed, in shocked regions, the difference

$e_{\text{tot}} - e_{\text{kin}}$ contains the adiabatic change of e_{th} and e_{nt} and the kinetic energy converted into thermal energy by the shock. In the absence of a subgrid model for nonthermal energy generation within shocks, the conservative choice is to assume that all non-adiabatic energy increase across shocks is thermalized,

$$e_{\text{th}} = e_{\text{tot}} - \frac{\rho \mathbf{u}^2}{2} - e_{\text{nt}}, \quad (3)$$

while the energies of the nonthermal components change adiabatically.

In general, nonthermal energies can be generated within shocks, e.g., via cosmic ray acceleration and turbulence driving, which can be taken into account in the partitioning of $e_{\text{tot}} - e_{\text{kin}}$ between e_{th} and e_{nt} (Section 2.3). It is also worth noting that the shock structure and jump conditions are generally expected to be modified by kinetic non-MHD effects and propagation of cosmic rays across the shock (e.g., Voelk et al. 1984; Haggerty & Caprioli 2020; Bret 2020; Hin Navin Tsung et al. 2020). These modifications are uncertain and depend on the details and physical parameters of shocks, which generally are not possible to model self-consistently in cosmological simulations. Thus, the effects of such modifications need to be modeled phenomenologically as a part of the CR shock injection model.

In the remainder of this section, we outline two methods implemented in the ART code for solving Equations (1)–(2) and our choices of synchronization criteria that determine when e_{th} is reset using Equation (3). These methods were originally developed for modeling thermal energy in highly supersonic flows by Bryan et al. (1995) and Ryu et al. (1993), and in this paper we explore their behavior in the context of modeling nonthermal energies. In principle, different energy components can be treated using different methods, however, we opt for a consistent treatment of all components using the same method.

2.1. Energy-based scheme

In the “dual energy” formulation proposed by Bryan et al. (1995), Equations (1)–(2) are solved in their original form: the energies e_{th} and e_{nt} are advected using conservative mass fluxes, and the PdV work is added as a source term, making the scheme non-conservative.

In Godunov-type methods, these source terms can be computed relatively accurately and efficiently because the time-averaged velocity of gas at a given interface is computed during the solution of the Riemann problem across cell interfaces and can be used for an accurate estimate of $\nabla \cdot \mathbf{u}$.

This method of advancing e_{th} and e_{nt} is used in regions where the flow is highly supersonic and the values

² For multiple nonthermal components, $e_{\text{nt},i}$, each component i obeys Equation (2) and its individual equation of state ($P_{\text{nt},i} = (\gamma_{\text{nt},i} - 1)e_{\text{nt},i}$, in our implementation); the total pressure of gas becomes $P = P_{\text{th}} + \sum_i P_{\text{nt},i}$, and e_{nt} in the expression for e_{tot} and Equation (3) is replaced by $\sum_i e_{\text{nt},i}$.

of e_{th} from e_{tot} can become smaller than the truncation error of the scheme, meaning highly inaccurate. In regions of modest Mach number, e_{th} is synchronized with e_{tot} using Equation (3). Specifically, the energies are synchronized in cells with

$$\frac{e_{\text{tot}} - e_{\text{kin}}}{e_{\text{tot}}} > \eta_1, \quad (4)$$

with the value of $\eta_1 = 10^{-3}$ suggested by Bryan et al. (1995). This condition is equivalent to the Mach number threshold of $M < \sqrt{2(\eta_1^{-1} - 1)/(\gamma(\gamma - 1))} \approx 40$.³ Additionally, Bryan et al. (1995) applies a second criterion based on the Mach numbers of adjacent cells:

$$\frac{e_{\text{tot},j} - e_{\text{kin},j}}{\max(e_{\text{tot},j-1}, e_{\text{tot},j}, e_{\text{tot},j+1})} > \eta_2, \quad (5)$$

where j denotes the cell to which the criterion is applied and $j \pm 1$ its immediate neighbors. We use the value of $\eta_2 = 0.1$ as in the original scheme of Bryan et al. (1995), which corresponds to $M \lesssim 4$ and makes this a significantly more conservative criterion than Equation (4).

Overall, these two criteria are used to select the method by which e_{th} is evolved—i.e., either by updating its value from e_{tot} (Equation (3)), or by solving Equation (1) independently of e_{tot} . We choose to always follow e_{nt} using Equation (2). For smaller η_1 and η_2 , e_{th} is evolved synchronously with e_{tot} in a larger fraction of the simulated volume, which in the limiting case of $\eta_1 = \eta_2 = 0$ corresponds to not using the dual energy formulation at all.

Note that our implementation slightly differs from the original scheme of Bryan et al. (1995), in which the first criterion is only used to identify the cells in which thermal pressure is computed using explicitly advected e_{th} , while only the second criterion is used to decide whether to synchronize e_{th} and e_{tot} . This is done to limit the dynamical effect of explicitly advected e_{th} on the solution. However, given that the first criterion selects only the gas with $e_{\text{th}} \ll e_{\text{tot}}$, the modeling of e_{th} in such gas should have negligible effect on the dynamics of the gas. We thus opt to use the first criterion to synchronize e_{th} with e_{tot} and always compute pressures from explicitly advected e_{th} and e_{nt} .

2.2. Entropy-conserving scheme

Ryu et al. (1993) proposed an alternative method, where a modified entropy, $\rho S_i \equiv P_i/\rho^{\gamma_i-1}$, is followed

³ In the absence of nonthermal energy components, this relation follows from $e_{\text{tot}} = e_{\text{kin}} + e_{\text{th}}$ and $e_{\text{kin}} = \rho u^2/2 = M^2 \rho c_s^2/2 = M^2 \gamma(\gamma-1)e_{\text{th}}/2$. In presence of nonthermal components, M and γ correspond to effective values.

as a separate variable instead of energy. The equations for ρS_i can be derived by combining Equations (1)–(2) with the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \quad (6)$$

which gives

$$\frac{\partial \rho S_{\text{th}}}{\partial t} + \nabla \cdot (\rho S_{\text{th}} \mathbf{u}) = 0, \quad (7)$$

$$\frac{\partial \rho S_{\text{nt}}}{\partial t} + \nabla \cdot (\rho S_{\text{nt}} \mathbf{u}) = 0. \quad (8)$$

These equations are in the conservative form and thus can be solved with a numerical scheme that conserves ρS_{th} and ρS_{nt} to machine precision.⁴ Given that the gas density can also be conserved to machine precision, this means that we can get accurate estimates of the entropy of the thermal and nonthermal energy components, S_{th} and S_{nt} , using ρS and ρ and computing the entropy fluxes, $S \rho \mathbf{u}$, consistently with the mass fluxes, $\rho \mathbf{u}$. Specifically, to compute the advection flux of entropy across a given cell interface, we interpolate the value of $S_i = P_i/\rho^{\gamma_i}$ (which can be thought of as a proxy for entropy per unit mass) at the upwind side of the interface using the same reconstruction scheme as for all other variables, and multiply it by the average mass flux from the solution of the Riemann problem (Springel 2010).⁵

The criteria for synchronizing S_i with the total energy of the gas, e_{tot} , also differ qualitatively from those in the energy-based method. Indeed, thermal entropy can be generated in shocks and Equation (7) therefore becomes invalid in regions with strong shocks, meaning that e_{tot} must be used instead. Thus, apart from the threshold on the Mach number of the flow, additional criteria must be used to identify shocked regions. For example, Ryu et al. (1993) check if the flow is locally converging ($\nabla \cdot \mathbf{u} < 0$) and the pressure jump across the cell is larger than a chosen threshold value.

An alternative method to identify shocked regions was proposed by Springel (2010, Section 3.5). This method is based on the idea that in Godunov-type methods the production of entropy is taken care of by the shock(s)

⁴ In fact, as was pointed out by Ryu et al. (1993), one can construct an arbitrary number of conservative quantities of the form $P^\alpha \rho^{1-\gamma\alpha}$ with $\alpha \in (-\infty, \infty)$. This can be easily checked by expanding the conservation equation for such a quantity and using Equations (1), (2), and (6). The entropy conservation form corresponds to $\alpha = 1$.

⁵ We also tried using the “raw” value of S_i from the upwind cell without reconstruction but found that using reconstructed values of S_i better preserves sharp contact discontinuities in different pressure components.

present in the Riemann solution for the hydrodynamic fluxes across cell interfaces. For weak shocks, the dissipated energy is strongly subdominant to the adiabatic increase of internal energy, and therefore one can identify the regions with strong production of entropy by defining a threshold in the Mach number of such “Riemann shocks.” Springel (2010) suggests the value $M_{R,\text{crit}} = 1.1$, so that e_{tot} is used in the regions where $M_R > M_{R,\text{crit}}$ on at least one of the cell’s interfaces. Although the relation between the shocks in the Riemann solutions and the physical shocks in the simulated flow is not direct, the total entropy generated by the real shock accumulates from the increments produced by the “Riemann shocks” on the interfaces resolving the real shock, which warrants the above criterion.

In our experiments, we find that the $M_{R,\text{crit}}$ threshold alone cannot prevent spurious heating in regions with strong velocity gradients. Such gradients can be interpreted by the Riemann solver as a discontinuity producing a sufficiently strong shock with a Mach number that can satisfy the above criterion. To filter out such cases, we require that the velocity of the shock(s) in the Riemann solution must be sufficiently large compared to the local flow velocity. In addition, to avoid injection of entropy in poorly resolved highly supersonic flows, we also require that the right-hand side of Equation (3) is not strongly subdominant to e_{tot} , which is analogous to the criteria used by Ryu et al. (1993) and Bryan et al. (1995).

Based on extensive experiments we find that the following set of criteria work well to identify shock regions:

$$\max(M_R) > 1.1, \quad (9)$$

$$|s_{R,\text{max}}|/|v_{\text{cell}}| > 0.1, \quad (10)$$

$$(e_{\text{tot}} - e_{\text{kin}} - e_{\text{nt}})/e_{\text{tot}} > 10^{-3}, \quad (11)$$

where $\max(M_R)$ is the maximal Mach number of the “Riemann shocks” at the interfaces of a given cell, $|s_{R,\text{max}}|$ is the corresponding shock velocity, and $|v_{\text{cell}}|$ is the gas velocity in the cell. When all of the above criteria are satisfied, the thermal energy is synchronized with total energy according to Equation (3), meaning that thermal entropy is injected into the cell in the amount corresponding to the energy dissipated by the shock. Otherwise, the evolution of thermal energy is followed using the entropy conservation Equation (7). To enforce adiabatic behavior of nonthermal energy components, their evolution is always followed by Equation (8).

As was pointed out by Springel (2010), one significant disadvantage of such a scheme is that it forfeits strict conservation of energy. In structure and galaxy formation simulations, however, the total energy is usually not conserved anyway owing to radiative cooling, radiative

or feedback heating, star formation, etc. All of these processes are uncertain and modeled approximately. At the same time, as we show below, in the idealized tests, the entropy-conserving scheme performs either comparably or better than the energy-based scheme. If the strict conservation of energy is nevertheless desirable, one may use the energy-based scheme for thermal energy, computing e_{th} from e_{tot} when possible, and the entropy-conserving scheme for all nonthermal components.

2.3. Generation of nonthermal energies in shocks

Synchronizing e_{th} with e_{tot} according to Equation (3) implies that all kinetic energy dissipated by shocks is thermalized. However, the synchronization procedure can easily be modified to account for nonthermal energy generation on shocks, such as cosmic ray acceleration or driving of small-scale turbulence when these components are modeled as e_{nt} . Indeed, if the adiabatic evolution of e_{th} and e_{nt} can be enforced during a hydrodynamic step, then at the end of the step the difference $e_{\text{diss}} \equiv e_{\text{tot}} - e_{\text{kin}} - e_{\text{th}} - e_{\text{nt}}$ in each cell will correspond to the total energy dissipated by shocks during the step. Therefore, to convert a fraction ζ of this energy into a nonthermal component, one only needs to add ζe_{diss} to e_{nt} and the remaining $(1 - \zeta)e_{\text{diss}}$ to e_{th} . For $\zeta = 0$, this scheme is equivalent to using Equation (3).

As we demonstrate in Section 3.2, this scheme works remarkably well when the adiabatic index of the injected energy is the same as that of the thermal energy, and it produces reasonable results when the adiabatic indices are different (e.g., for cosmic rays with $\gamma_{\text{nt}} = 4/3$). At the same time, this scheme is trivial to implement within the entropy-conserving scheme as it requires only a minor modification of the energy synchronization algorithm, while a number of previous implementations of energy injection on shocks required on-the-fly shock-finding algorithms. Note that if ζ depends on the properties of the shock, as is generally the case for cosmic rays, the shock-finding algorithm is still needed to measure these properties for physical shocks. For example, one may try to estimate the local properties of the shock by considering several adjacent cells or using more complex algorithms (e.g., Ryu et al. 2003; Pfrommer et al. 2006; Skillman et al. 2008; Schaal & Springel 2015). However, the energy partitioning scheme can still be used to inject the corresponding nonthermal energy.

3. IDEALIZED TESTS

3.1. Shock tube test with adiabatic nonthermal energy

Figure 1 shows the results of a shock tube test with an additional nonthermal fluid representing cosmic rays

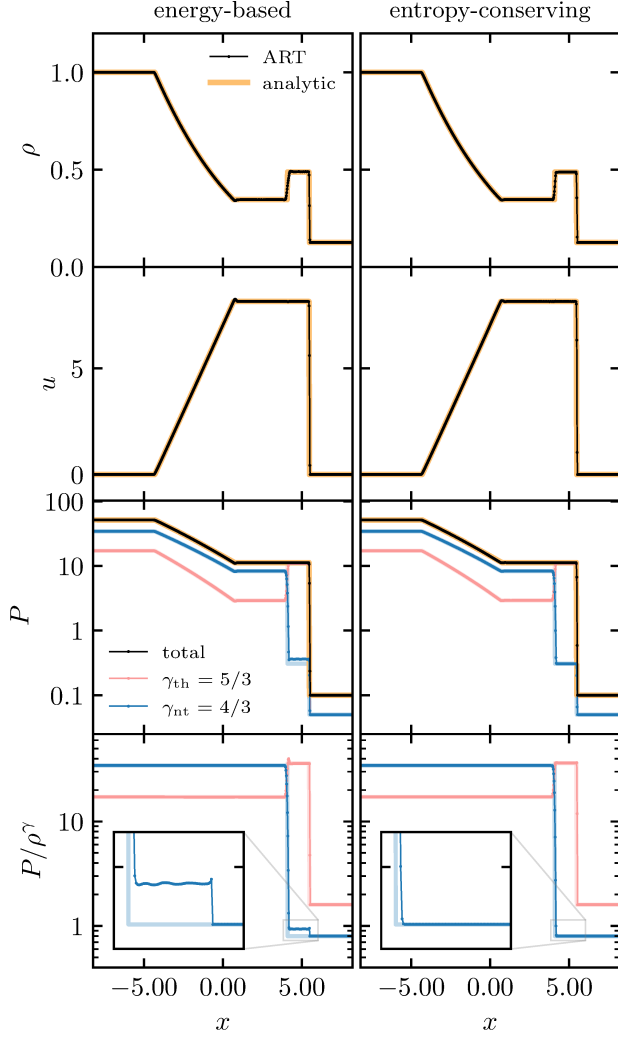


Figure 1. Shock tube test with an extra nonthermal energy representing cosmic rays ($\gamma_{nt} = 4/3$) using the energy-based (left column) and entropy-conserving (right column) schemes for modeling thermal energy and cosmic rays. The initial conditions are taken from Pfrommer et al. (2017): $(\rho, u, P_{th}, P_{nt}) = (1, 0, 17.172, 34.344)$ and $(0.125, 0, 0.05, 0.05)$ for the left and right initial states, respectively. Thick lines show the analytic solution from Pfrommer et al. (2017, Appendix C2), with orange lines showing the solution for density, velocity, and total pressure, and red and blue lines showing the solution for thermal and cosmic ray components, respectively. The results are compared at $t = 0.25$. As the bottom panel shows, the energy-based scheme leads to a spurious production of nonthermal entropy at the shock, while the entropy-conserving scheme ensures nonthermal entropy conservation across the shock.

(CRs) with $\gamma_{nt} = 4/3$ (Pfrommer et al. 2006, 2017). The setup is analogous to the classic Sod (1978) test and has the same structure of the solution except that pressure consists of two components with different adia-

batic indices, meaning that the effective γ can vary from region to region. As the figure shows, both energy-based and entropy-conserving methods reproduce the analytic solution for density, velocity and total pressure, which demonstrates that the fluxes of mass, momentum, and total energy are computed correctly in both schemes.

The difference between the methods appears in the solution for CR entropy and pressure as highlighted in the inset panels in the bottom row. In the energy-based scheme, CR entropy is spuriously produced at the shock, leading to a $\sim 20\%$ excess of the post-shock CR pressure. The post-shock thermal pressure becomes underestimated by the same (absolute) amount because in the post-shock region e_{th} is computed from e_{tot} and e_{cr} using Equation (3). In contrast, the entropy-conserving scheme does not suffer from such errors: CR entropy is exactly conserved across the shock, resulting in correct CR and thermal pressures in the post-shock region.

The spurious production of nonthermal entropy in the energy-based scheme stems from the fact that inviscid Equations (1) and (2) are formally invalid at shocks because of divergent gradients. In a numerical solution, however, shocks are smeared over several cells via numerical diffusivity and all gradients are finite. As the gas flows through such shocked regions, the PdV source term computed from the local pressure and velocity gradients mixes the adiabatic compression with a fraction of the energy dissipated by the shock, leading to a non-adiabatic behavior. In the absence of nonthermal energies, this effect does not introduce severe errors because, in the post-shock region, the solution for the advected e_{th} is usually disregarded and e_{th} is reset from the difference $e_{tot} - e_{kin}$. In contrast, if nonthermal energies are present, only the sum $e_{th} + e_{nt}$ can be set to $e_{tot} - e_{kin}$ while the individual values of e_{th} and e_{nt} become affected by the entropy error.

The $\sim 20\%$ error in the above test may seem small and unlikely to be important in real applications, e.g., galaxy formation simulations, due to much more uncertain source terms such as those involved in star formation and feedback modeling. However, we find that the error strongly increases for more compressible shocks and larger adiabatic indices of the nonthermal component. Indeed, as the entropy error originates from the PdV source term ($P_{nt} \nabla \cdot \mathbf{u} = P_{nt} \partial u / \partial x$ for a 1-dimensional problem), its magnitude relative to e_{nt} increases with γ_{nt} because $P_{nt} = (\gamma_{nt} - 1)e_{nt}$. The magnitude of the PdV term also increases with the compression ratio of the shock ($r_s = \rho_1 / \rho_0$, where “0” and “1” denote pre- and post-shock regions, respectively) because the magnitude of the velocity gradient increases

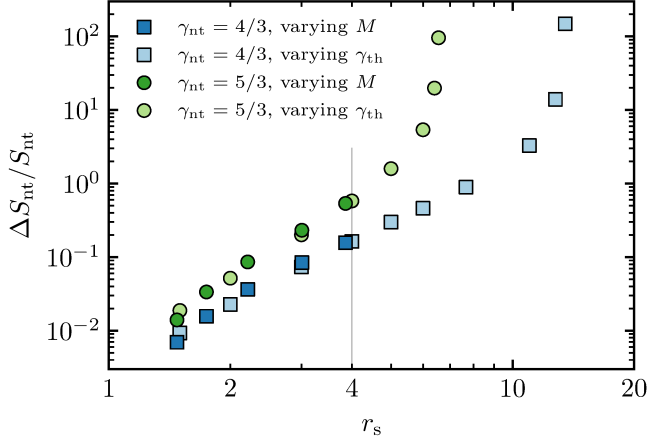


Figure 2. Dependence of the nonthermal entropy error produced by the energy-based scheme in the post-shock region on the compression ratio of the shock, $r_s = \rho_1/\rho_0$, for different values of the adiabatic index of the nonthermal component: $\gamma_{nt} = 4/3$ (blue squares) and $\gamma_{nt} = 5/3$ (green circles). The entropy error increases monotonically with r_s until $\Delta S_{nt}/S_{nt}$ becomes ~ 1 , at which point the error diverges. For a larger γ_{nt} , this divergence occurs at a smaller r_s . For this test, the nonthermal energy is dynamically decoupled from the hydrodynamics by making its initial value arbitrary small and removing its contribution from the total pressure. The color intensity corresponds to different ways of varying r_s . Lighter colors correspond to variations of γ_{th} , keeping the shock Mach number high so that $r_s \approx (\gamma_{th} + 1)/(\gamma_{th} - 1)$, with γ_{th} spanning a range from $\gamma_{th} = 5$ (leftmost points) to $\gamma_{th} = 1.16$ and 1.36 (rightmost points for $\gamma_{nt} = 4/3$ and $\gamma_{nt} = 5/3$, respectively). Darker colors show the variation of the shock Mach number at fixed $\gamma_{th} = 5/3$ (from left to right, $M \approx 1.3, 1.5, 1.9, 3.0, 9.0$), with the vertical gray line showing the high-Mach number limit of $r_s = 4$.

with r_s , and the errors are amplified by stronger compression.

Figure 2 shows the nonthermal entropy error in the post-shock region as a function of the compression ratio r_s for $\gamma_{nt} = 4/3$ (blue points) and $\gamma_{nt} = 5/3$ (green points). To exclude the dynamical effect of the error on the solution we set the initial value of e_{nt} to an arbitrary small value ($< 10^{-11}$ of e_{th} in the post-shock region) and rerun the shock tube test with varying the shock Mach number (i.e., varying the initial pressure jump) and adiabatic indices γ_{th} and γ_{nt} . Dark colors show the variation of the Mach number at fiducial $\gamma_{th} = 5/3$, which leads to variation of the compression ratio with the limiting value of $r_s = (\gamma_{th} + 1)/(\gamma_{th} - 1) = 4$ (shown with the vertical gray line). To explore variations of r_s beyond this limiting value, we also rerun the tests with different γ_{th} . Such a variation is relevant for practical applications, e.g., in galaxy formation simulations, be-

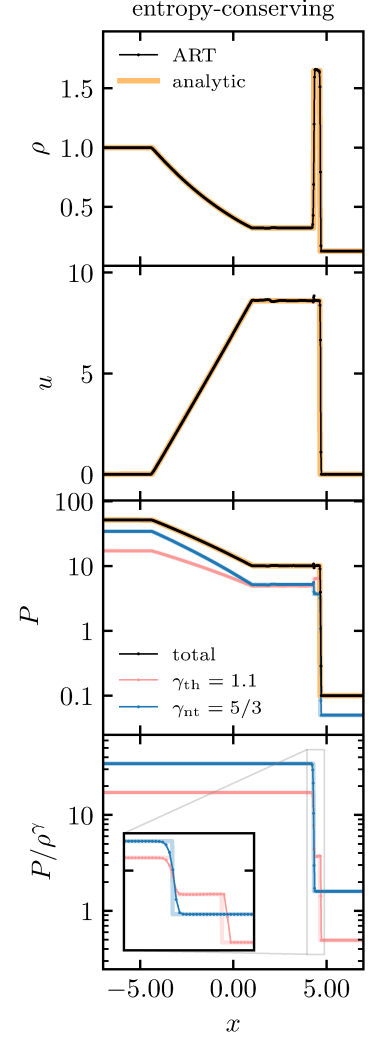


Figure 3. Results of the entropy-conserving scheme in the regime where the energy-based scheme is highly unstable. The initial conditions are the same as in Figure 1, but the adiabatic indices of the thermal and nonthermal components are set to $\gamma_{th} = 1.1$ and $\gamma_{nt} = 5/3$ to make the shock highly compressible, with a compression ratio of $r_s \approx 13.2$. Such r_s is significantly higher than the value at which the entropy error of the energy-based scheme diverges ($r_s \sim 6.5$ for $\gamma_{nt} = 5/3$, see Figure 2). The entropy-conserving scheme still performs well.

cause $r_s > 4$ can be achieved due to strong radiative cooling, even if γ_{th} is formally fixed at $5/3$.

The figure shows that while the entropy errors of the $\gamma_{nt} = 4/3$ component on strong shocks with $r_s \approx 4$ are $\sim 20\%$, the errors of the $\gamma_{nt} = 5/3$ component increase by a factor of 4. Thus, the spurious production of nonthermal entropy becomes more problematic for modeling components such as small-scale ISM turbulence with $\gamma = 5/3$ (e.g., Robertson & Goldreich 2012; Schmidt et al. 2014). Moreover, the error quickly increases as the

shock becomes more compressive, and this increase saturates only when the post-shock e_{nt} becomes comparable to the total energy dissipated by the shock, independent of the pre-shock value of e_{nt} . As the figure shows, for larger γ_{nt} , this divergence of errors occurs at smaller r_s : $r_s \sim 14$ and 6.5 for $\gamma_{\text{nt}} = 4/3$ and $5/3$, respectively. As we will show in Section 4, such errors can significantly change the behavior of nonthermal components in galaxy formation simulations.

In the entropy-conserving scheme, the nonthermal entropy conservation equation is followed in the entire simulation domain. Figure 3 demonstrates that this scheme can produce an adequate solution even for strongly compressive shocks for which the energy-based scheme is highly unstable.

3.2. Shock tube test with sourcing of nonthermal energy by the shock

The shock tube test in the previous section assumes that all energy dissipated by the shock is thermalized, so that nonthermal energy behaves adiabatically. In this section, we switch to the case where nonthermal energy is also sourced by the shock as described in Section 2.3. This method requires an accurate modeling of the adiabatic compression of both e_{th} and e_{nt} . The shock-generated entropy errors discussed in the previous section prevent the application of this method with the energy-based scheme; therefore, we only present the results for the entropy-conserving scheme.

Figure 4 shows the results of the test where 20% of the energy dissipated by the shock is converted into the nonthermal component, which is tracked separately and is shown with green lines. Different columns show the results for different values of the adiabatic index of the nonthermal component, namely, $\gamma_{\text{nt}} = 4/3$ and $5/3$. As this figure shows, for $\gamma_{\text{nt}} = 4/3$ the method underproduces nonthermal pressure and entropy in the post-shock region by a small amount, while for $\gamma_{\text{nt}} = 5/3$ the method works remarkably well.

Qualitatively, this difference can be attributed to the effect that the injection of e_{nt} changes the compressibility of the shock and to the fact that a shock is resolved by several cells, each dissipating shock energy. When γ_{nt} is smaller (larger) than $\gamma_{\text{th}} = 5/3$, converting a fraction of dissipated energy to e_{nt} results in a lower (higher) effective adiabatic index of gas in the post-shock region, making the shock more (less) compressible. The total change in the injected e_{nt} accumulates across the region over which the shock is numerically smeared and thus the effective adiabatic index also changes cell-by-cell gradually changing gas compressibility. As the left column in Figure 4 shows, the total change of e_{nt} does

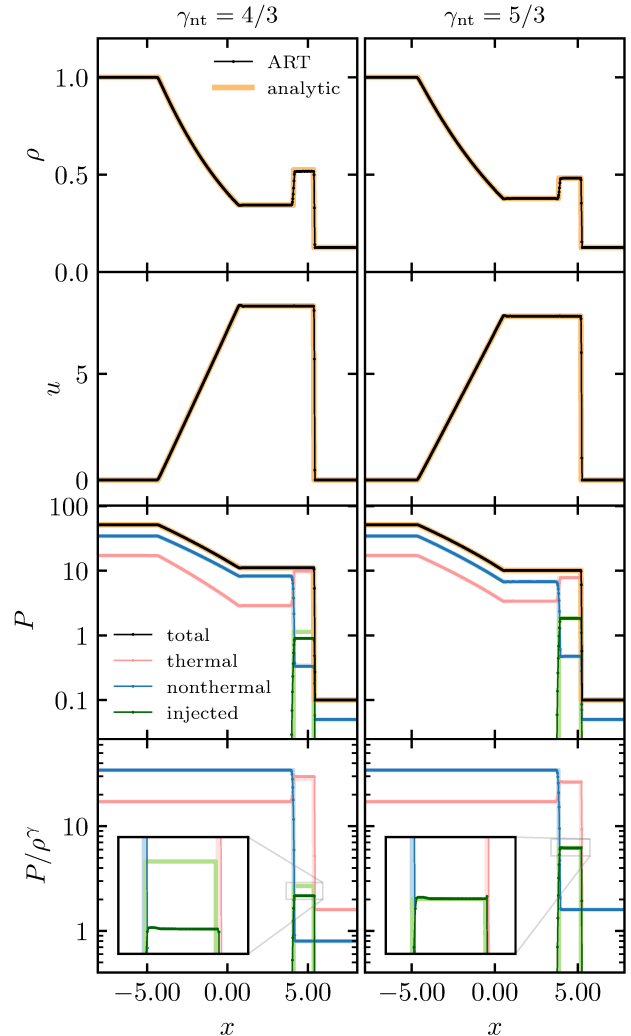


Figure 4. The shock tube test with 20% of dissipated energy injected in the nonthermal component using the entropy-conserving scheme and the energy partitioning method outlined in Section 2.3. The initial conditions are the same as in Figure 1. The left and right columns show the results for different adiabatic indices of the nonthermal component: $\gamma_{\text{nt}} = 4/3$ and $5/3$, respectively. The nonthermal energy injected at the shock is tracked separately and shown with green color in the bottom two panels. The analytic solution for all quantities is computed following Pfrommer et al. (2017, Appendix C2) and is shown with the thick lines. When γ_{nt} differs from $\gamma_{\text{th}} = 5/3$, the injected energy in the post-shock region is underestimated by $\sim 20\%$, while for $\gamma_{\text{nt}} = \gamma_{\text{th}} = 5/3$ the result is in perfect agreement with the analytical solution.

not add up correctly and results in smaller than expected injected e_{nt} for $e_{\text{nt}} = 4/3 < \gamma_{\text{th}}$. We also checked that $\gamma_{\text{nt}} > \gamma_{\text{th}}$ leads to an overestimate of the injected e_{nt} .

In contrast, when $\gamma_{\text{nt}} = \gamma_{\text{th}}$, the effective adiabatic index stays constant and therefore the injection of γ_{nt}

does not affect the compressibility of the gas. As the right column in the figure shows, our injection scheme works well in this case.

3.3. Zel'dovich pancake test

The formation of a Zel'dovich pancake is a stringent test for an implementation of thermal and nonthermal energies because it involves both a purely adiabatic stage with extremely supersonic gas motion and the formation of strong shocks after the crossing time.

The solution for density, $\rho(x)$, and velocity, $u(x)$, of a sine perturbation evolving from an initial redshift z_i to some later redshift z can be expressed in a parametric form (Zel'dovich 1970):

$$x(q, z) = q - \frac{1 + z_c}{1 + z} \frac{\sin(kq)}{k}, \quad (12)$$

$$\rho(q, z) = \rho_0 \left[1 - \frac{1 + z_c}{1 + z} \cos(kq) \right]^{-1}, \quad (13)$$

$$u(q, z) = -H_0 \frac{1 + z_c}{\sqrt{1 + z}} \frac{\sin(kq)}{k}, \quad (14)$$

where q is the Lagrangian coordinate, ρ_0 is the average density, $k = 2\pi/\lambda$ is the wavenumber of the initial perturbation, and the amplitude of the wave is described by the redshift of wave crossing, z_c .

For ease of comparisons, we choose parameters similar to Bryan et al. (1995), Trac & Pen (2004), and Springel (2010): $\lambda = 64h^{-1}$ Mpc and $z_c = 1$, and initialize the test at $z_i = 100$. We add a nonthermal energy component with $\gamma_{nt} = 5/3$ and initialize e_{th} and e_{nt} so that their initial entropies are constant with values corresponding to the average temperatures of $T_i = (\mu m_p/k) S_i \rho_0^{\gamma-1} = 30$ K and 70 K for thermal and nonthermal components, respectively (where μm_p is the average particle mass in units of the proton mass).

Both entropy components are expected to be conserved until the formation of the shocks at the wave crossing, and the solution for temperatures is therefore

$$T(q, z) = T_i \left[\left(\frac{1 + z}{1 + z_i} \right)^3 \frac{\rho(q, z)}{\rho_0} \right]^{\gamma-1}. \quad (15)$$

Figure 5 shows the results of the Zel'dovich test before ($z = 2$) and after ($z = 0$) the formation of shocks. Both energy-based and entropy-conserving schemes produce good results for gas density and velocity: during the adiabatic stage, e_{th} and e_{nt} are negligible and thus their treatment does not affect the solution, while the agreement after the wave crossing indicates that both methods capture the formation of physical shocks at $z = z_c = 1$.

In contrast, the solutions for the thermal and nonthermal components are quite different. As the lower left panel shows, before the crossing redshift, the energy-based scheme conserves nonthermal entropy to sub-percent level, while the thermal entropy suffers from strong numerical heating at the very center of the wave. The entropy-conserving scheme, in contrast, ensures entropy conservation for both components until the shock forms. After shock formation, both schemes produce similar results for thermal entropy, while the solution for nonthermal entropy in the energy-based scheme suffers from strong spurious heating at the shocks. The error has the same origin as the errors discussed in Section 3.1 and is completely absent in the entropy-conserving scheme.

The differences in thermal energy evolution before crossing illustrate the effects of the criteria selected to synchronize e_{th} and e_{nt} with $e_{tot} - e_{kin}$. Large value of thermal energy in the center in the energy-based scheme occurs because thermal energy is reset there from e_{tot} , while e_{th} is still smaller than e_{tot} by orders of magnitude. The truncation error of the scheme (typically $\sim 10^{-3}$ – $10^{-2} e_{tot}$) thus propagates into e_{th} as it is reset via $e_{th} = e_{tot} - e_{kin} - e_{nt}$, leading to orders of magnitude increase in the temperature at the wave center. This resetting in the center occurs because the criterion $(e_{tot} - e_{kin})/e_{tot} > \eta$ can be satisfied near the center of the wave, where $u \sim 0$ even though the flow is highly supersonic in the adjacent cells. Accounting for neighboring cells in the criterion (Equation (5)) or setting the threshold η to a larger value does not prevent this issue, but only delays its onset because the error in $e_{tot} - e_{kin}$ accumulates with time.

The entropy-conserving scheme adopts additional criteria based on the Mach number and velocity of the shocks in the Riemann solution (Equations (9) and (10)) that filter out such cases, ensuring entropy conservation until the wave crossing. In addition, given that during the adiabatic stage the conservation of entropy is enforced instead of e_{tot} , the errors in $e_{tot} - e_{kin}$ do not accumulate.

4. ISOLATED GALAXY SIMULATIONS

To test different methods for modeling nonthermal energies in a more realistic environment, we use simulations of an isolated $\sim L_*$ galaxy. Specifically, for the tests presented below, we use a snapshot from our fiducial simulation explored in Semenov et al. (2017, 2018, 2019) as the initial conditions. Below, we briefly describe the aspects of this simulation that are most relevant to the current study and refer the reader to our previous papers for more details.

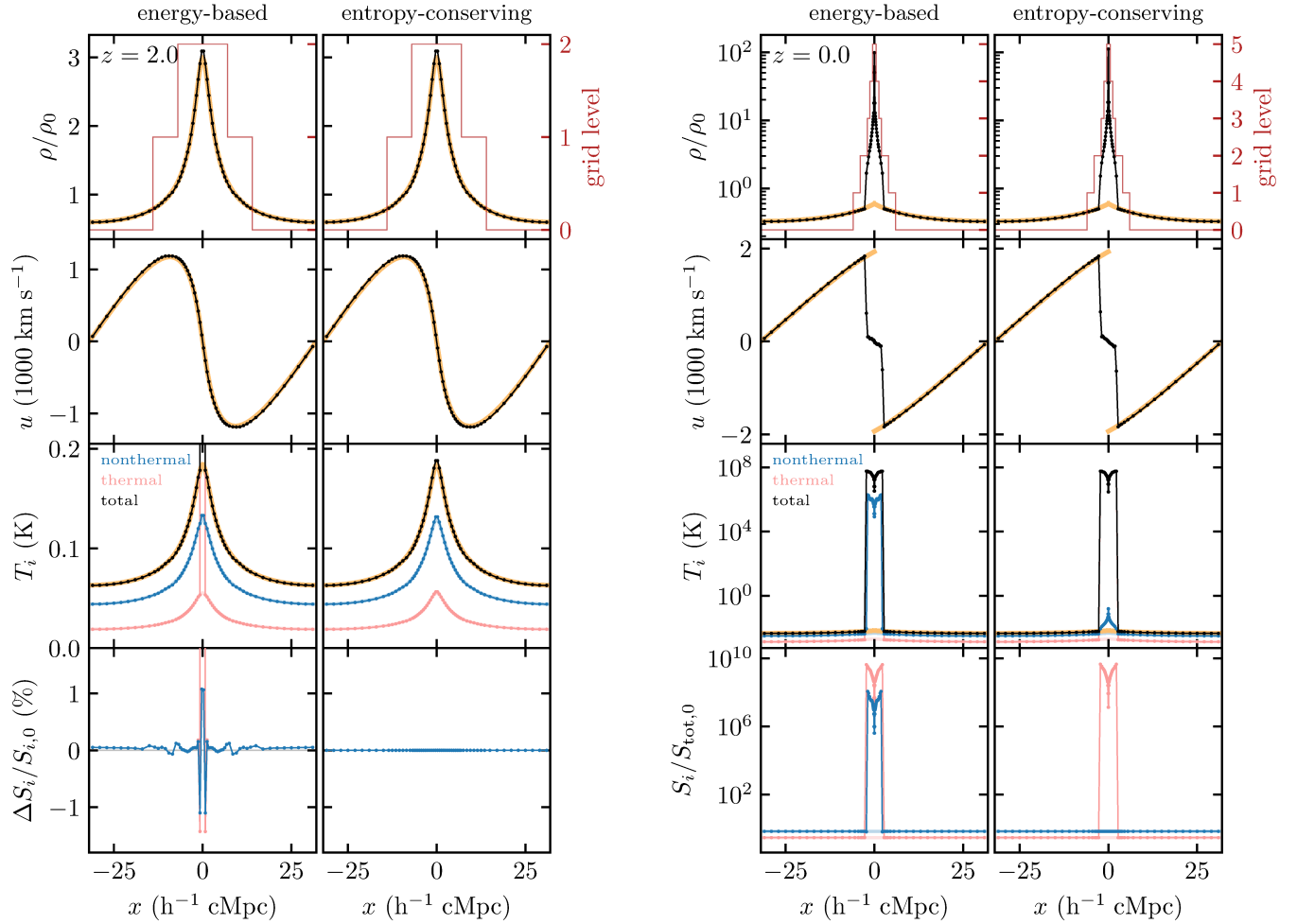


Figure 5. Hydrodynamic Zel’dovich pancake test with 70% of internal energy in the nonthermal form with $\gamma_{\text{nt}} = \gamma_{\text{th}} = 5/3$. The results are shown before and after the wave crossing (left and right sets of panels, respectively) and for the energy-based and entropy-conserving methods (left and right columns in each set). The connected points show simulation results while the thick lines show the analytic solution: orange lines show gas density, velocity, and total temperature, while red and blue lines show thermal and nonthermal components. Note that the analytic solution is valid only outside the shocked regions. Thin red lines in the top panels show the grid refinement level, with the 0-th level corresponding to 32 cells per box size. The bottom panels in the left set show the relative entropy error, while in the right set, they show the absolute value of entropy normalized to the initial total entropy. The energy-based scheme conserves nonthermal entropy relatively well during the adiabatic stage (with a relative error of $< 1\%$). After the wave crossing, however, the nonthermal entropy suffers from strong numerical heating at the shocks. The entropy-conserving scheme, in contrast, conserves nonthermal entropy throughout the duration of the run, with the nonthermal “temperature” increasing only due to adiabatic compression.

One of the key ingredients of these simulations is the dynamic model for unresolved turbulence, which is implemented as a nonthermal energy. Our implementation (Semenov et al. 2016) is based on the “shear-improved” model of Schmidt et al. (2014). In this model, the turbulent energy on unresolved scales, e_{turb} , is modeled as a nonthermal energy component with an adiabatic index of $\gamma = 5/3$, with source and sink terms that describe the turbulent cascade from the resolved velocity fluctuations and dissipation of turbulence on the local cell-crossing time. Apart from these source and sink terms, the sub-

grid turbulent energy is equivalent to e_{nt} as introduced and tested in Sections 2 and 3.

The initial conditions for our isolated galaxy simulations are taken from the AGORA code comparison project (Kim et al. 2014). The AMR grid cells are adaptively refined when the gas mass in a cell exceeds $\sim 8300 M_{\odot}$ until the minimal cell size of $\Delta = 40$ pc is reached. Radiative cooling and heating of gas are modeled using the Gnedin & Hollon (2012) model, assuming constant metallicity at the solar value and a constant radiation background with the H_2 photodissociation rate in the Lyman–Werner bands of 10^{-10} s^{-1} (Stecher &

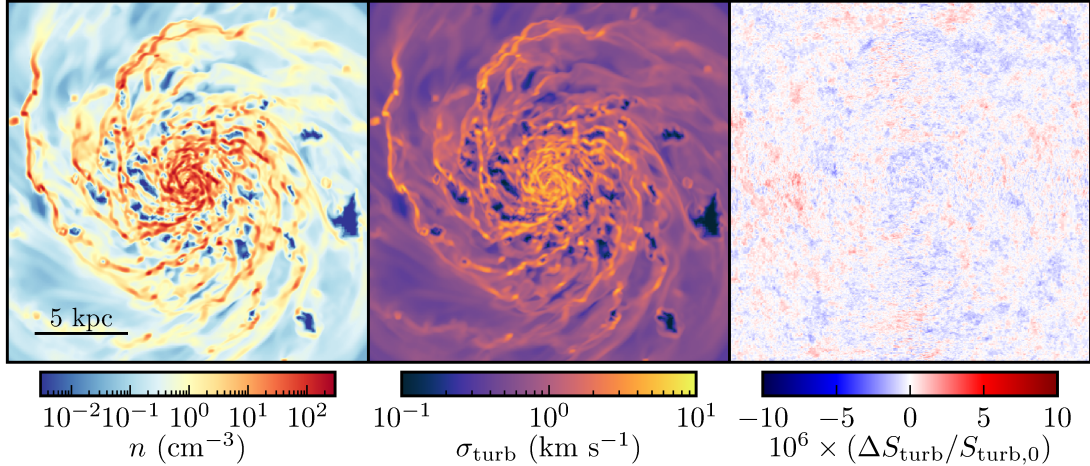


Figure 6. Gas density (left), velocity dispersion of subgrid turbulence (center) and the error of the subgrid turbulence entropy (right) in an isolated disk galaxy simulation after ~ 300 Myr of evolution. The subgrid turbulence is modeled using the entropy-conserved scheme. In this test, subgrid turbulence is initialized such that its entropy is initially constant, and then the galaxy simulation is evolved with all the source and sink terms of the subgrid turbulence model turned off, except for the PdV term. The right panel illustrates the long-term conservation of the entropy of subgrid turbulence, which remains constant to machine precision ($\sim 10^{-6}$ for a single-precision floating point variable) after ~ 300 Myr of evolution.

Williams 1967). We also adopt a model for dense gas shielding, which is calibrated against radiative transfer simulations of the ISM (the “L1a” model from Safranek-Shrader et al. 2017).

Local star formation is parametrized via the star formation efficiency per freefall time, $\dot{\rho}_\star = \epsilon_{\text{ff}}\rho/t_{\text{ff}}$, with a fixed value of $\epsilon_{\text{ff}} = 1\%$. Star-forming gas is defined by using a threshold either in density (Section 4.1) or in the local virial parameter (Section 4.2), where the latter is computed from the local turbulent velocity, $\sigma_{\text{turb}} = \sqrt{2e_{\text{turb}}/\rho}$, by using the definition from Bertoldi & McKee (1992):

$$\alpha_{\text{vir}} \equiv \frac{5\sigma_{\text{tot}}^2 R}{3GM} \approx 9.35 \frac{(\sigma_{\text{tot}}/10 \text{ km s}^{-1})^2}{(n/100 \text{ cm}^{-3})(\Delta/40 \text{ pc})^2}. \quad (16)$$

Stellar feedback is implemented by injecting radial momentum and thermal energy in the amounts calibrated against high-resolution simulations of SN remnants evolution in a nonuniform medium (Martizzi et al. 2015), with the momentum boosted by a factor of 5 to account for the effects of SN clustering (e.g., Gentry et al. 2017, 2018) and cosmic rays (Diesing & Caprioli 2018). To approximate the effects of pre-SN feedback, we start the injection of momentum from the moment of star particle formation and continue the injection at a constant rate for 40 Myr. The injection rate is set by the total number of SNe occurring for a given star particle, which we compute assuming the Chabrier (2003) IMF. We also account for stellar mass loss by using the Leitner & Kravtsov (2011) model.

4.1. Adiabatic subgrid turbulence energy

We start by demonstrating that, in a realistic ISM with a wide variation of density and temperature, our implementation of the entropy-conserving scheme ensures a long-term conservation of nonthermal entropy. To this end, we initialize the subgrid turbulent energy such that its entropy is constant and turn off all its source and sink terms. As such a model for turbulence is not realistic, we use a star formation threshold not in α_{vir} but in density, defining all gas with $n > n_{\text{sf}} = 100 \text{ cm}^{-3}$ as star-forming. We select this threshold value because, for our simulated galaxy, it results in roughly the same amount of star-forming gas as the fiducial α_{vir} threshold (Semenov et al. 2018). For this test, we also modified the original star formation and refinement algorithms such that the nonthermal entropy of gas in a cell is conserved when a star particle is formed or the cell is refined or defined.

Figure 6 shows the maps of gas density, subgrid turbulent velocity, and entropy error after 300 Myr of evolution. In the absence of source and sink terms, the nonthermal entropy maintains its initial value to machine precision, $\sim 10^{-6}$ for the single-precision floating point used in this test.

We do not show the results for the energy-based scheme because there the entropy errors discussed in Section 3.1 quickly accumulate resulting in a steady increase of turbulent pressure that eventually stabilizes the disk. In a simulation with a more realistic processes, such runaway heating does not occur because the cooling term in the equation for the nonthermal energy leads to

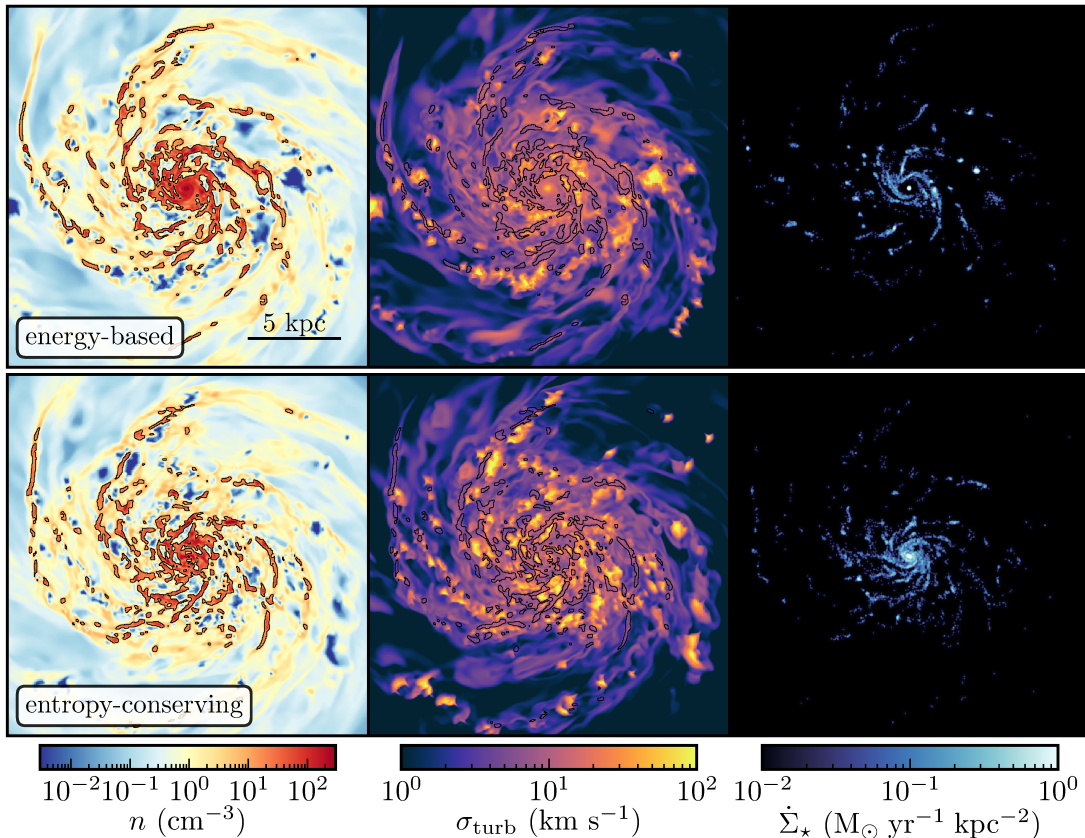


Figure 7. Results of the isolated galaxy simulation with the full subgrid turbulence model and star formation prescription coupled with the predicted turbulent velocities. The columns show the mid-plane slices of density, n , and subgrid turbulent velocity, σ_{turb} , and the surface density of SFR, $\dot{\Sigma}_*$, computed using particles younger than 30 Myr. Different rows show simulations where thermal energy and subgrid turbulence are modeled by using either the energy-based or entropy-conserving scheme. To guide the eye, black contours in the n and σ_{turb} maps show $n = 10 \text{ cm}^{-3}$ density isocontours. The $\dot{\Sigma}_*$ maps show the most apparent difference between the two schemes: the SFR in the run with the energy-based scheme is more clumpy and strongly suppressed near the galaxy center.

efficient dissipation of erroneously generated turbulent energy. Therefore, in the next subsection, we compare the schemes with all the sink and source terms activated.

4.2. Full subgrid turbulence model

As a more realistic test, we rerun our galaxy simulation including all source and sink terms in the subgrid turbulence model. We also use a star formation threshold based on the subgrid virial parameter, $\alpha_{\text{vir}} < \alpha_{\text{vir,sf}} = 10$, which was our fiducial choice in Semenov et al. (2017, 2018, 2019).

Figure 7 compares the maps of gas density, subgrid turbulent velocity, σ_{turb} , and SFR surface density in simulations with the energy-based and entropy-conserving schemes. Although the overall distributions of n and σ_{turb} are quite similar, the figure reveals interesting differences. With the entropy-conserving scheme, the overall density structure becomes more flocculent, especially near the disk center. The difference in the $\dot{\Sigma}_*$ map is even more striking: while with the energy-based

scheme the $\dot{\Sigma}_*$ distribution is rather clumpy and has a strong deficiency in the central $\sim 1 \text{ kpc}$, in the simulation with the entropy-conserving scheme, $\dot{\Sigma}_*$ is smoother and is not suppressed near the center.

To make the comparison more quantitative and explain these differences, Figure 8 shows joint distributions of gas in the n - σ_{tot} plane colored according to the galactocentric radius of the cells and with the thick dotted line showing the star formation threshold.

In the warm and diffuse part of the ISM with $n < 10 \text{ cm}^{-3}$, the overall shapes of these distributions are similar. In most of such gas, σ_{tot} is dominated by the sound speed, $c_s \sim 4\text{--}20 \text{ km s}^{-1}$, which sets the lower envelope of the distribution at $n < 10 \text{ cm}^{-3}$, reminiscent of the shape of the n - T diagram at such densities. We checked that the distributions of σ_{turb} alone are also quite similar at $n < 10 \text{ cm}^{-3}$ in both schemes.

In contrast, there are two interesting differences in the cold and dense ISM, $n > 10 \text{ cm}^{-3}$, where subgrid turbulence dominates over thermal energy. First, the

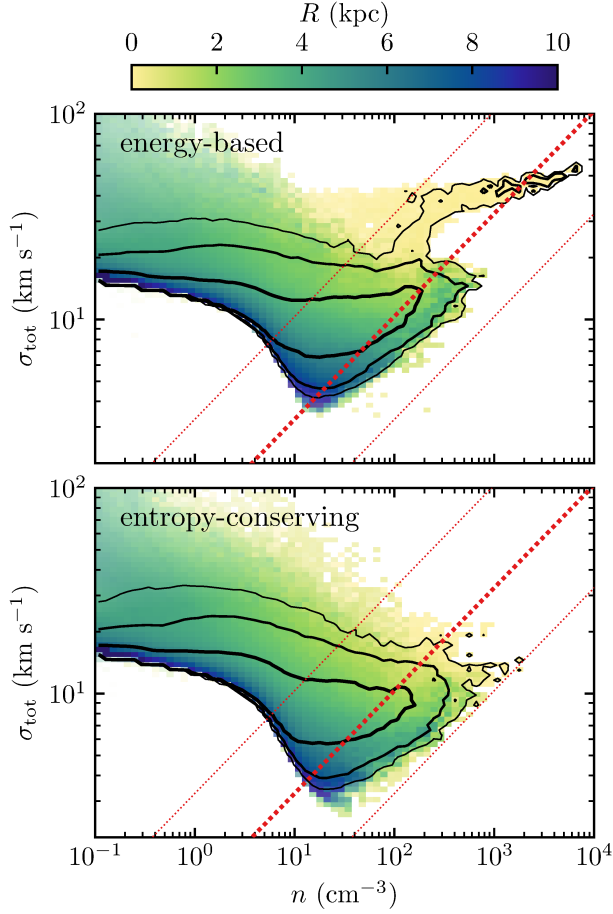


Figure 8. Joint distribution of gas density and total subgrid velocity dispersion, which includes both subgrid turbulence and thermal sound speed, $\sigma_{\text{tot}} = \sqrt{\sigma_{\text{turb}}^2 + c_s^2}$. The contours show 68%, 95%, and 99% of the gas mass and the color indicates the average galactocentric radius of cells within a pixel. The thick dotted line shows the star formation threshold, $\alpha_{\text{vir,sf}} = 10$, while the thin lines show the values of $\alpha_{\text{vir}} = 1$ and 100 for reference. To reduce noise, the distributions are averaged over 7 snapshots spaced 20 Myr apart. Switching to the entropy-conserving scheme reduces $\sigma_{\text{tot}} \approx \sigma_{\text{turb}}$ in dense gas ($n > 10 \text{ cm}^{-3}$) by a factor of ~ 1.5 and completely removes the horizontal feature at $\sigma_{\text{turb}} > 30 \text{ km s}^{-1}$ and $n > 100 \text{ cm}^{-3}$ that corresponds to the central region of the disk.

turbulent velocities in the main part of the distribution decrease by a factor of ~ 1.5 when the entropy-conserving scheme is used. Second, with the energy-based scheme, the gas near the disk center manifests itself as a prominent feature with $n > 100 \text{ cm}^{-3}$ and $\sigma_{\text{turb}} \sim 30\text{--}50 \text{ km s}^{-1}$, while with the entropy-conserving scheme, this feature disappears, and σ_{turb} in the central region becomes a continuation of the rest of the distribution.

The source of these differences is the production of e_{turb} in radiative shocks via the mechanism described in Section 3.1. The average distribution of cold gas in the $n\text{--}\sigma_{\text{tot}}$ plane is set by the competition between turbulence production and its decay on the local cell-crossing time. While both simulations include physical production terms—adiabatic compression and sourcing by fluctuating velocities—the numerical heating discussed in Section 3.1 also contributes to production of e_{turb} when the energy-based scheme is used. The fact that in most of the cold ISM the increase of σ_{turb} is only moderate indicates that this heating mechanism does not dominate over the physical production terms. The only exception is the disk center, where switching to the entropy-conserving scheme leads to a decrease of σ_{turb} by a factor of ~ 3 , which corresponds to a decrease of e_{turb} by an order of magnitude.

Even though in the context of our subgrid turbulence model the excess of e_{turb} in the disk center is purely numerical, it may reflect a physical enhancement of turbulence production in that region due to strong shear. Such an enhancement may not be captured by the adopted turbulence production term in the model, but it does appear as an enhancement of numerical dissipation of kinetic energy that is partially converted to e_{turb} (see Section 5.2 for further discussion). Therefore, here we solely present the difference between the two schemes of modeling nonthermal energies and leave the detailed investigation of the behavior in the disk center to a future study.

The decrease of σ_{turb} in the cold and dense gas leads to significant changes in the SFR distribution. For example, as σ_{turb} drops near the disk center, gas can cross the star formation threshold more easily in the run with the entropy-conserving scheme in striking contrast with the strongly suppressed SFR in the central $\sim 1 \text{ kpc}$ in the run with the energy-based scheme (recall the right panels in Figure 7). At larger galactocentric radii, distribution of star-forming regions also becomes different. As Figure 8 shows, when the entropy-conserving scheme is used, a larger fraction of gas from the main part of the distribution (i.e., excluding the feature corresponding to the disk center) reaches low virial parameters, $\alpha_{\text{vir}} < \alpha_{\text{vir,sf}} = 10$, and the typical densities of such star-forming gas become lower.

Figure 9 quantifies the effects of the turbulent energy scheme on the SFR, total mass of star-forming gas, M_{sf} , and its average density, $\bar{n}_{\text{sf}}$, at $R > 1 \text{ kpc}$. These three

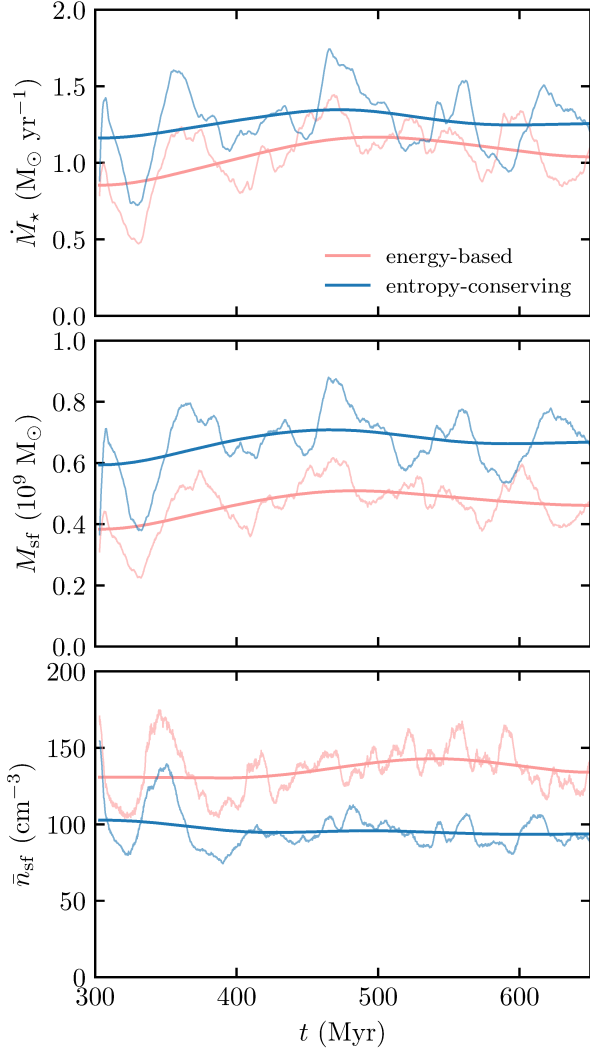


Figure 9. The effect of thermal and nonthermal energy modeling on the star formation in the average disk, i.e., at $R > 1$ kpc. The panels show the total star formation rate, mass of star-forming gas, and its average density. When the entropy-conserving scheme is used, a lower σ_{turb} in dense gas (see Figure 8) results in a larger amount of star-forming gas but also in a lower average density of such gas. These trends partially cancel out, leading to only a moderate increase in the SFR.

quantities are related via

$$\begin{aligned} \dot{M}_* &= \int \dot{\rho}_* dV = \int_{\text{sf}} \frac{\epsilon_{\text{ff}} \rho}{t_{\text{ff}}} dV = \\ &= \epsilon_{\text{ff}} M_{\text{sf}} \left\langle \frac{1}{t_{\text{ff}}} \right\rangle_{\text{sf}} \propto \epsilon_{\text{ff}} M_{\text{sf}} \bar{n}_{\text{sf}}^{0.5}, \end{aligned} \quad (17)$$

where $\epsilon_{\text{ff}} = 1\%$ and the average density of star-forming gas is defined from $\langle 1/t_{\text{ff}} \rangle_{\text{sf}}^{-1} \equiv \sqrt{3\pi/(32G\mu m_{\text{p}} \bar{n}_{\text{sf}})}$, assuming $\mu = 1$ and $\langle \dots \rangle_{\text{sf}}$ denoting mass-weighted averages over star-forming regions.

As the figure shows, the effect of the scheme on these quantities is quite significant, at the level of $\sim 10\%$ – 30% . In the entropy-conserving scheme, n_{sf} is smaller and M_{sf} is larger, which is consistent with the change of the PDF in Figure 8. These changes partially cancel in the total SFR, as larger mass of star-forming gas is partly offset by the lower density and longer free-fall time of star-forming regions. The total SFR in the run with the entropy-conserving scheme thus increases only moderately.

The direction of these trends is consistent with the model for the origin of global gas depletion times, $\tau_{\text{dep}} \equiv M_{\text{g}}/\dot{M}_*$, and SFR from continuous and rapid gas cycling between actively star-forming and diffuse, non-star-forming states in the ISM (Semenov et al. 2017). In this model, the global depletion time is the sum of the times that gas spends in each of these states, which can be written as $\tau_{\text{dep}} = N_{\text{c}} t_{\text{nsf}} + \tau_{\text{ff}}/\epsilon_{\text{ff}}$, where N_{c} is the total number of cycles required for gas depletion, t_{nsf} is the average residence time of gas in the non-star-forming state, and $\tau_{\text{ff}}/\epsilon_{\text{ff}}$ is the average depletion time of the star-forming gas. Switching to the entropy-conserving scheme reduces the production of e_{turb} , making it easier for gas to cross the $\alpha_{\text{vir}} \propto \sigma_{\text{turb}}^2/n$ threshold, which leads to shorter t_{nsf} and longer τ_{ff} . These trends counteract in the expression for τ_{dep} , but they add up in the expression for the star-forming gas mass fraction: $f_{\text{sf}} = (\tau_{\text{ff}}/\epsilon_{\text{ff}})/\tau_{\text{dep}} = (\epsilon_{\text{ff}} N_{\text{c}} t_{\text{nsf}}/\tau_{\text{ff}} + 1)^{-1}$. As a result, the increase of $\dot{M}_* = M_{\text{g}}/\tau_{\text{dep}}$ is smaller than the increase of $M_{\text{sf}} = f_{\text{sf}} M_{\text{g}}$ as M_{g} stays roughly constant over the considered time interval.

5. DISCUSSION

5.1. Implications for galaxy formation simulations

The results presented in the previous sections show that the choice of the method to model thermal and nonthermal energy components matters. In particular, the energy-based scheme can lead to a spurious generation of both thermal and nonthermal energy components. This was apparent in the Zel’dovich pancake collapse test, which showed significant spurious temperature increase at the center of the pancake right before crossing. Similarly, Villasenor et al. (2020) showed that the mean IGM temperature in their simulations is sensitive to how thermal energy of the gas is modeled and to the parameters of the dual energy scheme (see their Section 3.4 and Figures 5 and 6).

This difference arose because shock heating of gas in collapsed regions of virialized halos could be reduced or suppressed by the condition for synchronizing thermal energy followed independently with the value computed from the total energy. The results of Villasenor et al.

(2020) indicate that these issues can be corrected by the proper choice of thermal energy synchronization parameters, meaning that both energy- and entropy-based schemes can accurately and robustly model thermal energy. However, our results show that the choice of such scheme for any *nonthermal* component(s), such as subgrid turbulence or cosmic rays, does make a difference. This is because nonthermal components are necessarily evolved separately from the total energy in the entire simulation domain, while such treatment of the thermal energy alone is only needed in a limited volume corresponding to highly-supersonic flows.

Many modern fluid dynamics codes employ different schemes to separately model the coherent bulk motion of the fluid and small-scale motions (Trac & Pen 2004; Springel 2010; Duffell & MacFadyen 2011, 2015; Hopkins 2015; Duffell 2016). Such “moving mesh” approaches alleviate the inaccuracy of the thermal energy calculation from the total energy by absorbing a significant fraction of kinetic energy into the motion of the mesh. However, our results are equally relevant for these approaches because they indicate that the entropy-conserving scheme is the preferred choice for modeling the energies of non-thermal components. Indeed, for the energy-based formulation, the errors in the nonthermal entropy originate from the nonadiabatic part of the explicit PdV source term inside numerically smeared shocks which occur both in Eulerian and moving-mesh fluid dynamics schemes.

How much the choice of the scheme matters depends on the problem at hand. In the relatively dense parts of the ISM, where cooling is efficient, the spurious generation of energy near shocks may be inconsequential as it is dissipated efficiently. On the other hand, as we showed in the previous section, this choice may result in non-negligible changes in the star formation rate of a galaxy, the structure of the ISM, and in particular qualitatively different star formation in the central regions of the simulated galaxy.

Regardless of numerical issues, a strong argument for the entropy-based scheme for nonthermal energy components can be made in models that include the injection or dissipation of energy in shocks. For example, both turbulence and cosmic ray energy can be generated at shock fronts and, as we discussed in Section 2.3, such generation can be handled with ease in the entropy-conserving scheme.

This approach can be especially advantageous for sourcing nonthermal energies on radiative shocks. The existing methods based on on-the-fly shock finding algorithms estimate dissipated energy by measuring the pressure jump across the shock. This can be inaccurate

when the pressure in the post-shock region is subject to significant radiative losses. This issue can be alleviated in the proposed method because the energy dissipated by shocks at all cell interfaces can be converted to non-thermal components before radiative losses occur. We leave a more detailed investigation of this issue and a comparison of different methods to future study.

Finally, the entropy-conserving scheme can provide another significant advantage for modeling subgrid turbulence because the resolved turbulent energy dissipated by numerical viscosity can be captured into the subgrid turbulent energy *implicitly*, as we outline in the next section. This allows one to avoid explicit modeling via source terms that require calibration and can be rather uncertain.

5.2. Sourcing subgrid turbulence via capturing kinetic energy dissipated by numerical viscosity

Explicit modeling of unresolved turbulence in galaxy formation simulations can significantly improve the modeling of processes that depend on the small-scale turbulent structure of the ISM, such as star formation, mixing and transport of metals, and the effect of gas clumping on cooling and chemical reaction rates. Such subgrid models of turbulence require a prescription for the energy transfer from the resolved flow to unresolved scales.

For example, in the Large Eddy Simulation (LES) approach adopted in Section 4.2 (see Garnier et al. 2009 and Sagaut 2006 for reviews), such source terms are modeled using explicit closure relations, which can be calibrated against direct high-resolution simulations of turbulence (e.g., Schmidt & Federrath 2011). These closures are the major source of uncertainty in modeling subgrid turbulence, especially when the onset of the turbulent cascade is not sufficiently resolved or the local flow is significantly different from the direct simulations used to calibrate the closures.

On the other hand, the strict enforcement of entropy conservation opens up a new, more generic way for modeling the energy cascade to unresolved scales via tracking the local dissipation of the resolved kinetic energy. Indeed, as our results show, this scheme can be used to accurately follow adiabatic evolution of thermal and nonthermal energies during each hydrodynamics step. In regions where thermal energy can be computed reliably via the difference of the total and kinetic energies, $e_{\text{tot}} - e_{\text{kin}}$, the change of this energy during a single step can be compared to the the adiabatically evolving sum $e_{\text{th}} + e_{\text{turb}}$ modeled separately using the entropy-

conserving scheme.⁶ The difference between the two estimates will then correspond to the total kinetic energy dissipated by numerical viscosity in each cell during the step.

In most galaxy simulations, this dissipated energy is converted directly into heat and then quickly radiates away. In the real ISM, however, the energy is transferred to turbulent motions on small scales and decays on a longer timescale, close to the local eddy-turnover time. The above method can be used to capture this energy and to convert it into heat on a physically-motivated time scale, e.g., turbulent cell-crossing time with an order-of-unity pre-factor that can be calibrated in direct simulations of turbulence as in LES simulations.

Even though the dissipation of kinetic energy in this method is purely numerical and happens on the scale of a few computational cells, empirical evidence suggests that the *rate* of dissipation is correct. Indeed, turbulent box simulations carried out with different codes generally find turbulent spectra consistent with theoretical expectations, suggesting that the rate of the energy transfer on resolved scales is modeled correctly (e.g., Kritsuk et al. 2011). Given that in developed turbulence, this energy transfer rate is equal to the dissipation rate on the viscous scale, the rate of numerical dissipation must also be equal to the physical dissipation rate.

One additional issue that needs to be considered in such an approach is that not all kinetic energy dissipated by numerical viscosity should source the subgrid turbulence. For example, numerical viscosity can dissipate energy in laminar shearing flows, but no turbulence should be generated in this case. In the context of the LES approach, this motivates usage of the “shear-improved” versions of the schemes, in which the bulk flow of the gas is factored out from the source terms of subgrid turbulence (L  v  que et al. 2007; Schmidt et al. 2014). Conversely, a fraction of the energy dissipated by resolved shocks can drive turbulent motions in the post-shock region instead of being fully thermalized. These examples show that an approach described above should include a model for how the dissipated energy is apportioned between thermal and subgrid turbulence energy based on the local flow properties, such as the shear-improved approach and the scheme outlined in Section 2.3.

The idea of unresolved turbulence sourcing by numerical dissipation is a part of a more general assumption upon which the concept of the Implicit Large Eddy Simulations (ILES) is based (see, e.g., Grinstein et al. 2007

for a review): as long as a hydrodynamics solver satisfies the basic physical properties of the equations of hydrodynamics, such as conservation of mass, momentum, and energy, as well as causality and positivity, the effect of truncation errors at the resolution scale is equivalent to the net effect of unresolved scales. In this context, the approach outlined above captures the energy transfer to unresolved scales for the explicitly modeled subgrid turbulence without explicit and uncertain terms and closure relations. Instead, the energy transfer is implicitly handled by the hydrodynamic solver. The subgrid turbulent energy can then be used to model star-formation, metal diffusion, and gas clumping factor in exactly the same way as in the LES simulations.

6. SUMMARY AND CONCLUSIONS

We have investigated the accuracy of energy-based and entropy-conservative schemes in modeling nonthermal energy components. These schemes were initially introduced for modeling thermal energies of highly supersonic flows where e_{th} is strongly subdominant to the total and kinetic energies and therefore cannot be accurately computed as the difference between the two. The energy-based scheme numerically solves the explicit equation for the energy density evolution, which cannot be written in conservative form. The entropy-conserving scheme, on the other hand, uses a conservative equation for modified entropy and can thus conserve the entropy of an energy component to machine precision.

We have examined the performance of the schemes in following the subgrid turbulence and cosmic ray energy in the standard shock tube and Zel’dovich pancake tests and simulations of a realistic isolated L_* galaxy. Our results can be summarized as follows.

1. The energy-based scheme results in a spurious generation of nonthermal energy on shocks in the shock tube and Zel’dovich pancake tests, while the entropy-conserving method evolves the energy adiabatically to machine precision (see Figures 1, 3, and 5).
2. The magnitude of the nonthermal entropy error in the post-shock region increases with the adiabatic index of the nonthermal component and with the compression ratio of the shock (see Figure 2). The latter is particularly relevant for the cold ISM where shocks can be strongly compressible due to efficient radiative cooling.
3. In simulations of an isolated L_* galaxy with a turbulence-based star formation prescription, switching between the energy-based and entropy-conserved schemes result in a qualitative change of

⁶ Note that in the presence of cooling, the dissipation of energy can be accounted for in such a comparison during a single step.

morphology, in particular, large qualitative difference of star formation in the center of the galaxy (Figure 7), and ≈ 20 –30% change of the star formation rate away from the center (Figure 9).

4. We outline and test a simple, physical method for the injection of nonthermal energy on shocks. This method can be used in conjunction with the entropy-conserving scheme and can simplify the implementation of, e.g., injection of cosmic rays and driving of subgrid turbulence by shocks (see Section 2.3 and Figure 4).

Finally, we discuss how the entropy-conserving scheme can provide a straightforward way to capture the kinetic energy dissipated by numerical viscosity into the subgrid turbulent energy *implicitly*, without explicit source terms that require calibration and can be rather uncertain. It will be interesting to investigate the performance of such an approach using idealized simulations of turbulence and realistic galaxy formation simulations.

Although both methods have been shown to perform well in modeling the evolution of thermal energy, our results indicate that the entropy-conserving scheme is a preferred choice for modeling nonthermal energy com-

ponents. This conclusion is equally relevant for Eulerian and moving-mesh fluid dynamics codes.

ACKNOWLEDGMENTS

We would like to thank Nick Gnedin, Peng Oh, and Tsun Hin Navin Tsung for the insightful discussion. Support for V.S. was provided by NASA through the NASA Hubble Fellowship grant HST-HF2-51445.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. A.K. was supported by the National Science Foundation grants AST-1714658 and AST-1911111 and NASA ATP grant 80NSSC20K0512. Analyses presented in this paper were greatly aided by the following free software packages: NumPy (van der Walt et al. 2011), SciPy (Jones et al. 2001–2016), Matplotlib (Hunter 2007), and yt (Turk et al. 2011). We have also used the Astrophysics Data Service (ADS) and arXiv preprint repository extensively during this project and the writing of the paper.

REFERENCES

- Agertz, O., & Kravtsov, A. V. 2015, *ApJ*, 804, 18, doi: [10.1088/0004-637X/804/1/18](https://doi.org/10.1088/0004-637X/804/1/18)
- . 2016, *ApJ*, 824, 79, doi: [10.3847/0004-637X/824/2/79](https://doi.org/10.3847/0004-637X/824/2/79)
- Agertz, O., Kravtsov, A. V., Leitner, S. N., & Gnedin, N. Y. 2013, *ApJ*, 770, 25, doi: [10.1088/0004-637X/770/1/25](https://doi.org/10.1088/0004-637X/770/1/25)
- Benincasa, S. M., Wadsley, J., Couchman, H. M. P., & Keller, B. W. 2016, *MNRAS*, 462, 3053, doi: [10.1093/mnras/stw1741](https://doi.org/10.1093/mnras/stw1741)
- Bertoldi, F., & McKee, C. F. 1992, *ApJ*, 395, 140, doi: [10.1086/171638](https://doi.org/10.1086/171638)
- Booth, C. M., Agertz, O., Kravtsov, A. V., & Gnedin, N. Y. 2013, *ApJL*, 777, L16, doi: [10.1088/2041-8205/777/1/L16](https://doi.org/10.1088/2041-8205/777/1/L16)
- Braun, H., & Schmidt, W. 2012, *MNRAS*, 421, 1838, doi: [10.1111/j.1365-2966.2011.19889.x](https://doi.org/10.1111/j.1365-2966.2011.19889.x)
- Braun, H., Schmidt, W., Niemeyer, J. C., & Almgren, A. S. 2014, *MNRAS*, 442, 3407, doi: [10.1093/mnras/stu1119](https://doi.org/10.1093/mnras/stu1119)
- Bret, A. 2020, *ApJ*, 900, 111, doi: [10.3847/1538-4357/aba68e](https://doi.org/10.3847/1538-4357/aba68e)
- Brook, C. B., Stinson, G., Gibson, B. K., Wadsley, J., & Quinn, T. 2012, *MNRAS*, 424, 1275, doi: [10.1111/j.1365-2966.2012.21306.x](https://doi.org/10.1111/j.1365-2966.2012.21306.x)
- Bryan, G. L., Norman, M. L., Stone, J. M., Cen, R., & Ostriker, J. P. 1995, *Computer Physics Communications*, 89, 149, doi: [10.1016/0010-4655\(94\)00191-4](https://doi.org/10.1016/0010-4655(94)00191-4)
- Buck, T., Pfrommer, C., Pakmor, R., Grand, R. J. J., & Springel, V. 2020, *MNRAS*, 497, 1712, doi: [10.1093/mnras/staa1960](https://doi.org/10.1093/mnras/staa1960)
- Chabrier, G. 2003, *PASP*, 115, 763, doi: [10.1086/376392](https://doi.org/10.1086/376392)
- Chan, T. K., Kereš, D., Hopkins, P. F., et al. 2019, *MNRAS*, 488, 3716, doi: [10.1093/mnras/stz1895](https://doi.org/10.1093/mnras/stz1895)
- Colella, P. 1985, *SIAM Journal on Scientific and Statistical Computing*, 6, 104, doi: [10.1137/0906009](https://doi.org/10.1137/0906009)
- Colella, P., & Glaz, H. M. 1985, *Journal of Computational Physics*, 59, 264, doi: [10.1016/0021-9991\(85\)90146-9](https://doi.org/10.1016/0021-9991(85)90146-9)
- Costa, T., Pakmor, R., & Springel, V. 2020, *MNRAS*, 497, 5229, doi: [10.1093/mnras/staa2321](https://doi.org/10.1093/mnras/staa2321)
- Diesing, R., & Caprioli, D. 2018, *PhRvL*, 121, 091101, doi: [10.1103/PhysRevLett.121.091101](https://doi.org/10.1103/PhysRevLett.121.091101)
- Duffell, P. C. 2016, *ApJS*, 226, 2, doi: [10.3847/0067-0049/226/1/2](https://doi.org/10.3847/0067-0049/226/1/2)
- Duffell, P. C., & MacFadyen, A. I. 2011, *ApJS*, 197, 15, doi: [10.1088/0067-0049/197/2/15](https://doi.org/10.1088/0067-0049/197/2/15)
- . 2015, *MNRAS*, 449, 2718, doi: [10.1093/mnras/stv459](https://doi.org/10.1093/mnras/stv459)
- Garnier, E., Adams, N., & Sagaut, P. 2009, *Large Eddy Simulation for Compressible Flows* (Springer Netherlands)
- Gentry, E. S., Krumholz, M. R., Dekel, A., & Madau, P. 2017, *MNRAS*, 465, 2471, doi: [10.1093/mnras/stw2746](https://doi.org/10.1093/mnras/stw2746)

- Gentry, E. S., Krumholz, M. R., Madau, P., & Lupi, A. 2018, ArXiv e-prints. <https://arxiv.org/abs/1802.06860>
- Gnedin, N. Y., & Hollon, N. 2012, ApJS, 202, 13, doi: [10.1088/0067-0049/202/2/13](https://doi.org/10.1088/0067-0049/202/2/13)
- Gnedin, N. Y., & Kravtsov, A. V. 2011, ApJ, 728, 88, doi: [10.1088/0004-637X/728/2/88](https://doi.org/10.1088/0004-637X/728/2/88)
- Governato, F., Brook, C., Mayer, L., et al. 2010, Nature, 463, 203, doi: [10.1038/nature08640](https://doi.org/10.1038/nature08640)
- Grinstein, F. F., Margolin, L. G., & Rider, W. J. 2007, Implicit Large Eddy Simulation: Computing Turbulent Fluid Dynamics, ed. F. F. Grinstein, L. G. Margolin, & W. J. Rider (Cambridge University Press)
- Haggerty, C. C., & Caprioli, D. 2020, ApJ, 905, 1, doi: [10.3847/1538-4357/abbe06](https://doi.org/10.3847/1538-4357/abbe06)
- Hin Navin Tsung, T., Oh, S. P., & Jiang, Y.-F. 2020, arXiv e-prints, arXiv:2008.10537. <https://arxiv.org/abs/2008.10537>
- Hopkins, P. F. 2015, MNRAS, 450, 53, doi: [10.1093/mnras/stv195](https://doi.org/10.1093/mnras/stv195)
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581, doi: [10.1093/mnras/stu1738](https://doi.org/10.1093/mnras/stu1738)
- Hopkins, P. F., Narayanan, D., & Murray, N. 2013, MNRAS, 432, 2647, doi: [10.1093/mnras/stt723](https://doi.org/10.1093/mnras/stt723)
- Hopkins, P. F., Quataert, E., & Murray, N. 2011, MNRAS, 417, 950, doi: [10.1111/j.1365-2966.2011.19306.x](https://doi.org/10.1111/j.1365-2966.2011.19306.x)
- . 2012, MNRAS, 421, 3488, doi: [10.1111/j.1365-2966.2012.20578.x](https://doi.org/10.1111/j.1365-2966.2012.20578.x)
- Hopkins, P. F., Wetzel, A., Keres, D., et al. 2017, ArXiv e-prints. <https://arxiv.org/abs/1707.07010>
- Hunter, J. D. 2007, CSE, 9, 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- Jefferson, S. M. R., Krumholz, M. R., Fujimoto, Y., et al. 2021, arXiv e-prints, arXiv:2105.11457. <https://arxiv.org/abs/2105.11457>
- Jones, E., Oliphant, T., Peterson, P., et al. 2001-2016, <http://www.scipy.org>
- Jubelgas, M., Springel, V., Enßlin, T., & Pfrommer, C. 2008, A&A, 481, 33, doi: [10.1051/0004-6361:20065295](https://doi.org/10.1051/0004-6361:20065295)
- Kim, J.-h., Abel, T., Agertz, O., et al. 2014, ApJS, 210, 14, doi: [10.1088/0067-0049/210/1/14](https://doi.org/10.1088/0067-0049/210/1/14)
- Kravtsov, A. V. 1999, PhD thesis, NEW MEXICO STATE UNIVERSITY
- Kravtsov, A. V., Klypin, A., & Hoffman, Y. 2002, ApJ, 571, 563, doi: [10.1086/340046](https://doi.org/10.1086/340046)
- Kretschmer, M., Dekel, A., & Teyssier, R. 2021, arXiv e-prints, arXiv:2103.06882. <https://arxiv.org/abs/2103.06882>
- Kretschmer, M., & Teyssier, R. 2020, MNRAS, 492, 1385, doi: [10.1093/mnras/stz3495](https://doi.org/10.1093/mnras/stz3495)
- Kritsuk, A. G., Nordlund, Å., Collins, D., et al. 2011, ApJ, 737, 13, doi: [10.1088/0004-637X/737/1/13](https://doi.org/10.1088/0004-637X/737/1/13)
- Leitner, S. N., & Kravtsov, A. V. 2011, ApJ, 734, 48, doi: [10.1088/0004-637X/734/1/48](https://doi.org/10.1088/0004-637X/734/1/48)
- Lévêque, E., Toschi, F., Shao, L., & Bertoglio, J.-P. 2007, Journal of Fluid Mechanics, 570, 491, doi: [10.1017/S0022112006003429](https://doi.org/10.1017/S0022112006003429)
- Marinacci, F., Sales, L. V., Vogelsberger, M., Torrey, P., & Springel, V. 2019, MNRAS, 489, 4233, doi: [10.1093/mnras/stz2391](https://doi.org/10.1093/mnras/stz2391)
- Martizzi, D., Faucher-Giguère, C.-A., & Quataert, E. 2015, MNRAS, 450, 504, doi: [10.1093/mnras/stv562](https://doi.org/10.1093/mnras/stv562)
- Naab, T., & Ostriker, J. P. 2017, ARA&A, 55, 59, doi: [10.1146/annurev-astro-081913-040019](https://doi.org/10.1146/annurev-astro-081913-040019)
- Pakmor, R., Pfrommer, C., Simpson, C. M., & Springel, V. 2016, ApJL, 824, L30, doi: [10.3847/2041-8205/824/2/L30](https://doi.org/10.3847/2041-8205/824/2/L30)
- Pfrommer, C., Pakmor, R., Schaal, K., Simpson, C. M., & Springel, V. 2017, MNRAS, 465, 4500, doi: [10.1093/mnras/stw2941](https://doi.org/10.1093/mnras/stw2941)
- Pfrommer, C., Springel, V., Enßlin, T. A., & Jubelgas, M. 2006, MNRAS, 367, 113, doi: [10.1111/j.1365-2966.2005.09953.x](https://doi.org/10.1111/j.1365-2966.2005.09953.x)
- Robertson, B., & Goldreich, P. 2012, ApJL, 750, L31, doi: [10.1088/2041-8205/750/2/L31](https://doi.org/10.1088/2041-8205/750/2/L31)
- Rudd, D. H., Zentner, A. R., & Kravtsov, A. V. 2008, ApJ, 672, 19, doi: [10.1086/523836](https://doi.org/10.1086/523836)
- Ruszkowski, M., Yang, H. Y. K., & Zweibel, E. 2017, ApJ, 834, 208, doi: [10.3847/1538-4357/834/2/208](https://doi.org/10.3847/1538-4357/834/2/208)
- Ryu, D., Kang, H., Hallman, E., & Jones, T. W. 2003, ApJ, 593, 599, doi: [10.1086/376723](https://doi.org/10.1086/376723)
- Ryu, D., Ostriker, J. P., Kang, H., & Cen, R. 1993, ApJ, 414, 1, doi: [10.1086/173051](https://doi.org/10.1086/173051)
- Safrank-Shrader, C., Krumholz, M. R., Kim, C.-G., et al. 2017, MNRAS, 465, 885, doi: [10.1093/mnras/stw2647](https://doi.org/10.1093/mnras/stw2647)
- Sagaut, P. 2006, Large Eddy Simulation for Incompressible Flows (Springer-Verlag Berlin Heidelberg)
- Salem, M., & Bryan, G. L. 2014, MNRAS, 437, 3312, doi: [10.1093/mnras/stt2121](https://doi.org/10.1093/mnras/stt2121)
- Salem, M., Bryan, G. L., & Hummels, C. 2014, ApJL, 797, L18, doi: [10.1088/2041-8205/797/2/L18](https://doi.org/10.1088/2041-8205/797/2/L18)
- Schaal, K., & Springel, V. 2015, MNRAS, 446, 3992, doi: [10.1093/mnras/stu2386](https://doi.org/10.1093/mnras/stu2386)
- Schmidt, W., & Federrath, C. 2011, A&A, 528, A106, doi: [10.1051/0004-6361/201015630](https://doi.org/10.1051/0004-6361/201015630)
- Schmidt, W., Almgren, A. S., Braun, H., et al. 2014, MNRAS, 440, 3051, doi: [10.1093/mnras/stu501](https://doi.org/10.1093/mnras/stu501)
- Semenov, V. A., Kravtsov, A. V., & Gnedin, N. Y. 2016, ApJ, 826, 200, doi: [10.3847/0004-637X/826/2/200](https://doi.org/10.3847/0004-637X/826/2/200)
- . 2017, ApJ, 845, 133, doi: [10.3847/1538-4357/aa8096](https://doi.org/10.3847/1538-4357/aa8096)
- . 2018, ApJ, 861, 4, doi: [10.3847/1538-4357/aac6eb](https://doi.org/10.3847/1538-4357/aac6eb)
- . 2019, ApJ, 870, 79, doi: [10.3847/1538-4357/aaf163](https://doi.org/10.3847/1538-4357/aaf163)

- . 2021, arXiv e-prints, arXiv:2103.13406.
<https://arxiv.org/abs/2103.13406>
- Skillman, S. W., O’Shea, B. W., Hallman, E. J., Burns, J. O., & Norman, M. L. 2008, *ApJ*, 689, 1063, doi: [10.1086/592496](https://doi.org/10.1086/592496)
- Smith, M. C., Bryan, G. L., Somerville, R. S., et al. 2020, arXiv e-prints, arXiv:2009.11309.
<https://arxiv.org/abs/2009.11309>
- Sod, G. A. 1978, *Journal of Computational Physics*, 27, 1, doi: [10.1016/0021-9991\(78\)90023-2](https://doi.org/10.1016/0021-9991(78)90023-2)
- Springel, V. 2000, *MNRAS*, 312, 859, doi: [10.1046/j.1365-8711.2000.03187.x](https://doi.org/10.1046/j.1365-8711.2000.03187.x)
- . 2010, *MNRAS*, 401, 791, doi: [10.1111/j.1365-2966.2009.15715.x](https://doi.org/10.1111/j.1365-2966.2009.15715.x)
- Springel, V., & Hernquist, L. 2003, *MNRAS*, 339, 289, doi: [10.1046/j.1365-8711.2003.06206.x](https://doi.org/10.1046/j.1365-8711.2003.06206.x)
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, 475, 676, doi: [10.1093/mnras/stx3304](https://doi.org/10.1093/mnras/stx3304)
- Stecher, T. P., & Williams, D. A. 1967, *ApJL*, 149, L29, doi: [10.1086/180047](https://doi.org/10.1086/180047)
- Stinson, G. S., Brook, C., Macciò, A. V., et al. 2013, *MNRAS*, 428, 129, doi: [10.1093/mnras/sts028](https://doi.org/10.1093/mnras/sts028)
- Teyssier, R. 2015, *ARA&A*, 53, 325, doi: [10.1146/annurev-astro-082214-122309](https://doi.org/10.1146/annurev-astro-082214-122309)
- Teyssier, R., Pontzen, A., Dubois, Y., & Read, J. I. 2013, *MNRAS*, 429, 3068, doi: [10.1093/mnras/sts563](https://doi.org/10.1093/mnras/sts563)
- Trac, H., & Pen, U.-L. 2004, *NewA*, 9, 443, doi: [10.1016/j.newast.2004.02.002](https://doi.org/10.1016/j.newast.2004.02.002)
- Turk, M. J., Smith, B. D., Oishi, J. S., et al. 2011, *ApJS*, 192, 9, doi: [10.1088/0067-0049/192/1/9](https://doi.org/10.1088/0067-0049/192/1/9)
- Uhlig, M., Pfrommer, C., Sharma, M., et al. 2012, *MNRAS*, 423, 2374, doi: [10.1111/j.1365-2966.2012.21045.x](https://doi.org/10.1111/j.1365-2966.2012.21045.x)
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, 13, 22, doi: [10.1109/MCSE.2011.37](https://doi.org/10.1109/MCSE.2011.37)
- van Leer, B. 1979, *Journal of Computational Physics*, 32, 101, doi: [10.1016/0021-9991\(79\)90145-1](https://doi.org/10.1016/0021-9991(79)90145-1)
- Villasenor, B., Robertson, B., Madau, P., & Schneider, E. 2020, arXiv e-prints, arXiv:2009.06652.
<https://arxiv.org/abs/2009.06652>
- Voelk, H. J., Drury, L. O., & McKenzie, J. F. 1984, *A&A*, 130, 19
- Vogelsberger, M., Genel, S., Sijacki, D., et al. 2013, *MNRAS*, 436, 3031, doi: [10.1093/mnras/stt1789](https://doi.org/10.1093/mnras/stt1789)
- Vogelsberger, M., Marinacci, F., Torrey, P., & Puchwein, E. 2020, *Nature Reviews Physics*, 2, 42, doi: [10.1038/s42254-019-0127-2](https://doi.org/10.1038/s42254-019-0127-2)
- Wiener, J., Pfrommer, C., & Oh, S. P. 2017, *MNRAS*, 467, 906, doi: [10.1093/mnras/stx127](https://doi.org/10.1093/mnras/stx127)
- Yepes, G., Kates, R., Khokhlov, A., & Klypin, A. 1997, *MNRAS*, 284, 235, doi: [10.1093/mnras/284.1.235](https://doi.org/10.1093/mnras/284.1.235)
- Zel’dovich, Y. B. 1970, *A&A*, 5, 84