Title

FIRE-2 simulations: physics versus numerics in galaxy formation

Journal

Monthly Notices of the Royal Astronomical Society, 480(1)

ISSN

0035-8711

Authors

Hopkins, Philip F
Wetzel, Andrew
Kereš, Dušan
et al.

Publication Date

2018-10-11

DOI

10.1093/mnras/sty1690

Peer reviewed

Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 22 February 2017 (MN LATEX style file v2.2)

FIRE-2 Simulations: Physics versus Numerics in Galaxy Formation

Philip F. Hopkins⋆[1], Andrew Wetzel[1,2,3]†, Dušan Kereš[4], Claude-André Faucher-Giguère[5], Eliot Quataert[6], Michael Boylan-Kolchin[7], Norman Murray[8], Christopher C. Hayward[9], Shea Garrison-Kimmel[1], Cameron Hummels[1], Robert Feldmann[6,10], Paul Torrey[11], Xiangcheng Ma[1], Daniel Anglés-Alcázar[5], Kung-Yi Su[1], Matthew Orr[1], Denise Schmitz[1], Ivanna Escala[1], Robyn Sanderson[1], Michael Y. Grudić[1], Zachary Hafen[5], Ji-Hoon Kim[12], Alex Fitts[7], James S. Bullock[13], Coral Wheeler[1], T. K. Chan[4], Oliver D. Elbert[13], Desika Narayanan[14]

[1]*TAPIR, Mailcode 350-17, California Institute of Technology, Pasadena, CA 91125, USA*
[2]*Carnegie Observatories, Pasadena, CA 91101, USA*
[3]*Department of Physics, University of California, Davis, CA 95616, USA*
[4]*Department of Physics, Center for Astrophysics and Space Science, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA*
[5]*Department of Physics and Astronomy and CIERA, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA*
[6]*Department of Astronomy and Theoretical Astrophysics Center, University of California Berkeley, Berkeley, CA 94720*
[7]*Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712, USA*
[8]*Canadian Institute for Theoretical Astrophysics, 60 St. George Street, University of Toronto, ON M5S 3H8, Canada*
[9]*Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA*
[10]*Institute for Computational Science, University of Zurich, Zurich CH-8057, Switzerland*
[11]*Department of Physics, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*
[12]*Kavli Institute for Particle Astrophysics and Cosmology, Department of Physics, Stanford University, Stanford, CA, USA*
[13]*Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*
[14]*Department of Astronomy, University of Florida, Gainesville, FL 32611, USA*

Submitted to MNRAS, February 2017

ABSTRACT

The Feedback In Realistic Environments (FIRE) project explores the role of feedback in cosmological simulations of galaxy formation. Previous FIRE simulations used an identical source code ("FIRE-1") for consistency. Now, motivated by the development of more accurate numerics – including hydrodynamic solvers, gravitational force softening, and supernova coupling algorithms – and the exploration of new physics (such as magnetic fields), we introduce "FIRE-2", an updated numerical implementation of FIRE physics for the GIZMO code. We run a suite of cosmological simulations and compare against FIRE-1 results: overall, FIRE-2 improvements do not qualitatively change galaxy-scale properties. We pursue an extensive study of numerics versus physics in galaxy simulations. The details of the star-formation algorithm, cooling physics, and chemistry have weak effects, provided that we include metal-line cooling and star formation occurs at higher-than-mean densities. We present several new resolution criteria for high-resolution galaxy simulations. Most galaxy-scale properties are remarkably robust to the numerics that we test, provided that: (1) Toomre masses (cold disk scale heights) are resolved; (2) the feedback coupling ensures conservation and isotropy, and (3) individual supernovae are time-resolved. As resolution increases, stellar masses and profiles converge first, followed by metal abundances and visual morphologies, followed by properties of winds and the circum-galactic medium (CGM). The central ($\sim$kpc) mass concentration of massive ($> L_*$) galaxies is most sensitive to numerics, in particular, how winds ejected into hot halos are trapped, mixed, and recycled into the galaxy. Multiple feedback mechanisms are required to reproduce observations: supernovae dominate on cosmological timescales, regulating stellar masses; stellar (OB/AGB) mass loss primarily fuels late-time star formation; radiative feedback helps to suppress instantaneous star formation in disks and accretion into low-mass galaxies. We provide tables, initial conditions, and the numerical algorithms that are required to reproduce our simulations.

Key words: galaxies: formation — galaxies: evolution — galaxies: active — stars: formation — cosmology: theory — methods: numerical

1 INTRODUCTION

Feedback from stars is an essential and still poorly-understood component of galaxy formation. In the absence of stellar feedback, most gas accreted into galaxies should cool rapidly on a timescale much shorter than the dynamical time, collapse, fragment, and turn into stars (Hopkins et al. 2011; Tasker 2011; Bournaud et al. 2010; Dobbs et al. 2011; Krumholz et al. 2011; Harper-Clark & Murray 2011). Cosmologically, efficient cooling inevitably results in most of the baryons turning into stars, producing galaxies much more massive than observed (Katz et al. 1996; Somerville & Primack 1999; Cole et al. 2000; Springel & Hernquist 2003b; Kereš et al.

⋆ E-mail:phopkins@caltech.edu
† Caltech-Carnegie Fellow

Table 1. Initial suite of FIRE-2 simulations run to redshift $z = 0$

Simulation Name	$M_{\rm halo}^{\rm vir}$ [$M_\odot$]	$R_{\rm vir}$ [kpc]	M_* [$M_\odot$]	$R_{1/2}$ [kpc]	$m_{i,1000}$ [$1000\,M_\odot$]	$\epsilon_{\rm gas}^{\rm MIN}$ [pc]	$r_{\rm DM}^{\rm conv}$ [pc]	Notes
Ultra-Faints ($M_{\rm halo} \lesssim 10^{10}\,M_\odot$):								
m09	2.4e9	35.6	9.4e3	0.29	0.25	1.1	65	early-forming, ultra-faint field dwarf
m10a	7.0e9	51.5	–	–	0.50	–	88	very late-forming, dark halo (no stars)
Low-Mass Dwarf ($M_{\rm halo} \sim 10^{10}\,M_\odot$) Survey:								
m10q	8.0e9	52.4	1.8e6	0.63	0.25	0.52	73	isolated dwarf, early-forming halo
m10g	8.0e9	52.4	5.3e6	0.89	0.50	2.0	110	early-forming
m10v	8.3e9	53.1	1.0e5	0.31	0.25	0.73	65	isolated dwarf, late-forming halo
m10d	8.5e9	53.6	1.3e6	0.50	0.50	2.0	77	intermediate-forming
m10c	9.0e9	54.6	4.7e5	0.30	0.50	2.0	92	late-forming
m10b	9.4e9	55.4	4.5e5	0.33	0.50	2.0	77	late-forming
m10e	1.0e10	57.1	1.8e6	0.58	0.50	2.0	120	late-forming
m10i	1.1e10	57.8	7.1e6	0.52	0.50	2.0	75	early-forming
m10l	1.1e10	57.8	1.2e7	0.72	0.50	2.0	110	early-forming
m10j	1.1e10	58.5	8.4e6	0.70	0.50	2.0	86	late-forming, dense environment
m10k	1.2e10	59.3	9.5e6	1.0	0.50	2.0	140	early-forming
m10m	1.2e10	59.4	1.3e7	0.89	0.50	2.0	120	early-forming
m10h	1.3e10	61.6	6.8e6	0.74	0.50	2.0	96	intermediate-forming
m10f	1.3e10	62.3	1.1e7	1.1	0.50	2.0	150	early-forming
m10y	1.4e10	63.9	1.0e7	0.96	0.26	0.21	74	early-forming, large core
Intermediate-Mass Dwarfs ($10^{10}\,M_\odot \lesssim M_{\rm halo} \lesssim 10^{12}\,M_\odot$):								
m10z	3.5e10	85.6	3.5e7	2.1	0.26	0.21	130	ultra-diffuse galaxy
m11a	4.1e10	90.5	1.2e8	2.7	2.1	4.3	310	diffuse, large core
m11b	4.3e10	92.2	1.1e8	2.4	2.1	2.9	250	intermediate-forming
m11q	1.4e11	136	4.1e8	2.7	0.88	0.71	120	early-forming, large core
m11c	1.4e11	138	8.1e8	2.7	2.1	0.40	250	intermediate-forming
m11v	3.2e11	177	2.4e9	2.5	7.0	1.3	310	multi-merger ongoing ($z = 0$)
m11f	5.0e11	208	2.4e10	2.6	12	0.9	280	quiescent late history
Milky Way-Mass "Latte" ($M_{\rm halo} \sim 10^{12}\,M_\odot$) Halos:								
m12i	1.2e12	275	6.5e10	2.9	7.0	0.38	150	"Latte" primary halo
m12f	1.6e12	306	8.0e10	4.0	7.0	0.51	130	MW-like halo
m12m	1.5e12	301	1.2e11	5.6	7.0	0.27	180	earlier-forming halo, boxy bulge
"Low"-Resolution Milky Way-Mass Halo Survey:								
m12i_LowRes	1.2e12	278	1.0e11	2.3	56	1.4	290	Low-resolution "Latte" halo
m12f_LowRes	1.6e12	310	1.3e11	3.1	56	1.4	310	Low-resolution MW-like halo
m12b_LowRes	1.4e12	291	9.8e10	1.5	56	1.4	300	early-forming halo
m12c_LowRes	1.3e12	285	9.2e10	1.6	56	1.4	310	late-forming halo
m12m_LowRes	1.5e12	302	1.4e11	5.0	56	1.4	360	early-forming halo
m12q_LowRes	1.6e12	308	1.2e11	1.9	56	1.4	240	early-forming halo
m12z_LowRes	8.7e11	251	4.3e10	6.0	33	8.0	520	little/no bulge, merger at $z \approx 0$
m12_ELVIS_Robin	1.6e12	306	6.7e10	3.4	56	1.5	400	late-forming, gas-rich in pair
m12_ELVIS_Batman	2.0e12	325	1.2e11	1.0	56	1.5	210	compact, early-forming in pair
m12_ELVIS_Thelma	1.1e12	272	7.0e10	3.6	32	2.0	260	MW-like in Local Group pair
m12_ELVIS_Louise	1.5e12	297	1.3e11	4.2	32	2.0	300	M31-like in Local Group pair
m12_ELVIS_Romeo	1.3e12	285	8.1e10	6.5	28	1.0	280	M31-like in Local Group pair
m12_ELVIS_Juliet	1.1e12	267	6.0e10	5.0	28	1.0	260	MW-like in Local Group pair

Parameters describing the initial suite of FIRE-2 simulations in this paper. Each simulation contains several (in some, several dozen) galaxies in the high-resolution zoom-in region; halo and stellar properties listed refer only to the main "target" halo around which the high-resolution zoom-in region is centered. All properties refer to our highest-resolution simulation using each initial condition. All units are physical.

(1) Name of simulation.

(2) $M_{\rm halo}^{\rm vir}$: Virial mass (following Bryan & Norman 1998) of the "target" halo at $z = 0$ in simulation with baryons. Each simulation can include a few, to a few dozen, independent halos and central galaxies (typically of lower mass) within the zoom-in region.

(3) $R_{\rm vir}$: Virial radius (spherical) of the main halo at $z = 0$.

(4) M_*: Stellar mass (within $\leq 3\,R_{1/2}$) of the central galaxy in the target halo at $z = 0$.

(5) $R_{1/2}$: Half-mass radius of stars in the central galaxy at $z = 0$ (see § 3).

(6) $m_{i,1000}$: the mass of baryonic (gas or star) particles, in units of $1000\,M_\odot$. Dark matter (DM) particle masses are $\approx 5\times$ larger, according to the universal baryon fraction.

(7) $\epsilon_{\rm gas}^{\rm MIN}$: Minimum gravitational force softening reached by gas in the simulation. For gas, gravitational force resolution is *identical* to hydrodynamic spatial resolution (the same spatial gas distribution appears in gravity and hydrodynamic equations at all times). The gravitational force softening, ϵ_i, therefore corresponds to the gas inter-particle separation, h_i: $\epsilon_i = h_i$. See § 4 for other definitions of "spatial resolution". Forces become exactly Keplerian (point-mass like) at $> 1.95\,\epsilon_{\rm gas}$ from a particle; the "Plummer-equivalent" softening is $\approx 0.7\,\epsilon_{\rm gas}$.

(8) $r_{\rm DM}^{\rm conv}$: Radius of convergence in DM properties, based on the Power et al. (2003) criterion, using our best estimate for where the DM-only simulations converge (see § 4). This is approximately the radius enclosing > 200 DM particles. We show below that convergence in DM profiles can in fact extend to much smaller radii in runs with baryons. § 4 shows that the DM force softening $\epsilon_{\rm DM}$ is much less important, as long as it is $\lesssim r_{\rm DM}^{\rm conv}$. In our default runs, $\epsilon_{\rm DM}$ is fixed with values $= 40$ pc for our dwarfs and $= 30$ pc in our MW-mass (m12) systems.

(9) Notes: Additional information on each simulation.

Table 2. Physics & Numerics Explored in This Paper (and Papers II & III)

Physics/Numerics	§	Effects in FIRE-2 Simulations	Guidelines
		Resolution:	
Mass Resolution	4.1	Most results converge after resolving the Toomre scale, some (e.g. massive galaxy morphology) depend on resolved winds/hot gas	Convergence criteria in § 4.1.3 (Eq. 5-7)
Collisionless (DM/Stellar) Force Softening	4.2	Irrelevant unless *extremely* small or very large values used, adaptive collisionless softenings require additional timestep limiters	Optimal range of values in § 4.2.3
Gas Force Softening	4.2	Forcing fixed softening generally has no effect, unless too large, then fragmentation & SF are artificially suppressed	Fully-adaptive softenings (matching gas) should be used
Timestep Criteria	4.3	Provided that standard stability criteria are met, this has no effect. Additional limiters needed for stellar evolution & adaptive softening	Standard limiters + Stellar (Eq. 12) + Adaptive softening (Eq. 13)
		(Magneto)-Hydrodynamics:	
Hydro Method (MFM vs. SPH)	5	Irrelevant for dwarfs. Important for massive galaxies with hot halos. SPH suppresses cooling & artificially allows clumpy winds to vent	Newer methods recommended
Artificial Pressure "Floors"	6	Unimportant unless set too large, then prevents real fragmentation. Double-counts "sub-grid" treatment of fragmentation with SF model	Do not use with self-gravity based SF models
Magnetic Fields, Conduction, Viscosity	F	Weak effects on sub-galactic scales (dense gas, morphology, turbulent ISM) (Not studied here, but in Su et al. 2016; effects in CGM could be larger)	See Su et al. (2016)
Metal Diffusion (sub-resolution mixing)	7.2 & F	Small effects on galaxy properties & dynamics, but potentially important for abundance distributions of stars	Best practice depends on numerical hydro method
		Cooling:	
Molecular Chemistry/Cooling	7 & B	No effect on galaxy properties or star formation (just a tracer). Not important star formation criterion if fragmentation is resolved	May be relevant at [Z/H] $\ll -3$, can be important for observational tracers
Low-Temperature Cooling ($T \ll 10^4$ K)	7 & B	Details have no dynamical effects because $t_{\rm cool} \ll t_{\rm dyn}$ in cold gas to opacity limit ($\sim 0.01\,M_\odot$). Relevant for observables in cold phase	*Some* needed to form cold clouds, details dynamically irrelevant
Metal-Line Cooling ($T \gtrsim 10^4$ K)	7 & B	Dominates cooling in metal-rich centers of "hot halos" around massive galaxies, and of individual SNe blastwaves	Needed: important in super-bubbles & "hot halos"
Photo-Heating (Background)	7 & B	Significantly suppresses star formation in small ($M_{\rm halo} \lesssim 10^{10}\,M_\odot$) dwarfs	Needed: dwarfs & CGM/IGM
		Star Formation:	
Self-Gravity (Virial) Criterion	8 & C	Negligible effect on galaxy properties (SF is feedback-regulated). More accurately identifies collapsing regions in high-dynamic range situations	Recommended see Appendix C for implementation
Density Threshold	8 & C	Negligible effect on galaxy properties (SF is feedback-regulated) Can be arbitrarily high with adaptive gas softenings	Should exceed galactic mean density; ideally, highest resolved densities
Jeans-Instability Criterion	8 & C	Negligible effect on galaxy properties (SF is feedback-regulated). Automatically satisfied in high-density, self-gravitating gas	Not necessary
Self-shielding/Molecular Criterion	8 & C	Negligible effect on galaxy properties (SF is feedback-regulated). Automatically satisfied in high-density, self-gravitating gas	Not necessary
"Efficiency" (Rate) at Resolution Limit	8 & C	Negligible effect on galaxy properties (SF is feedback-regulated). If artificially lowered, more dense gas "piles up" until same SFR achieved	$\sim 100\%$ per free-fall in *locally-self-gravitating* gas
		Stellar Feedback:	
Continuous Mass-Loss (OB & AGB)	9 & A	Primarily important as a late-time fuel source for SF Relatively weak "primary" feedback effects on galactic scales	Should couple as Appendix D Rates given in § A
Supernovae (Ia & II) ("How to Couple")	A & D Paper II	Type-II: Dominant FB mechanism on cosmological scales. Need to account for *PdV* work if Sedov phase un-resolved. Subgrid models should reproduce exact solutions, conserve mass, energy, & momentum, and converge	Should couple as § D Validation & convergence tests & criteria in Paper II
Radiative Feedback (Photo-Heating & Radiation Pressure)	A & E Paper III	"Smooths" SF in dwarfs (less bursty) & suppresses SF in dense gas. UV background dominates in dwarfs. Photo-electric heating unimportant. IR multiple-scattering effects weak, except in massive galaxy nuclei.	Need photo-heating & single-scattering rad. pressure (Paper III) Rad.-hydro algorithm sub-dominant

A cursory outline of the physics and numerics explored in this paper. All "standard" FIRE-2 simulations, including all in Table 1, are run with the *identical* simulation code and physics. However, to understand how physical and numerical changes influence our results, we systematically "turn off" different physics and vary the numerical method and/or resolution in the sections listed here.
(1) Physics/Numerics: what we consider.
(2) §: Section where we pursue a detailed study of the effects of each numerics/physics on galaxy formation.
(3) Effects in FIRE-2 Simulations: Overall summary of the effects of variation in the relevant physics or numerics, insofar as it is relevant (or not) for the predictions of our simulations. This applies only *for quantities discussed in this paper*, that is, global galaxy properties such as SFRs, stellar masses, sizes, and morphologies. For example, although we show that arbitrarily removing molecular chemistry from our cooling networks has no effect on predicted galaxy properties or star formation (because other cooling channels are available and molecular gas is primarily a tracer, not a causal driver of star formation), molecular chemistry is obviously fundamentally important if one wishes to predict molecular lines! Furthermore, we did not examine detailed properties of the CGM or IGM, where different physics may dominate.
(4) Guidelines: Approximate "rules of thumb" for the relevant physics or numerics in the context of "zoom-in" galaxy simulations. In the text, we provide more detailed guidelines. For example, for numerical resolution and other numerical parameters, we provide equations that approximately determine whether or not key physics should be resolved (and we show that these accurately represent when our predictions are numerically converged).

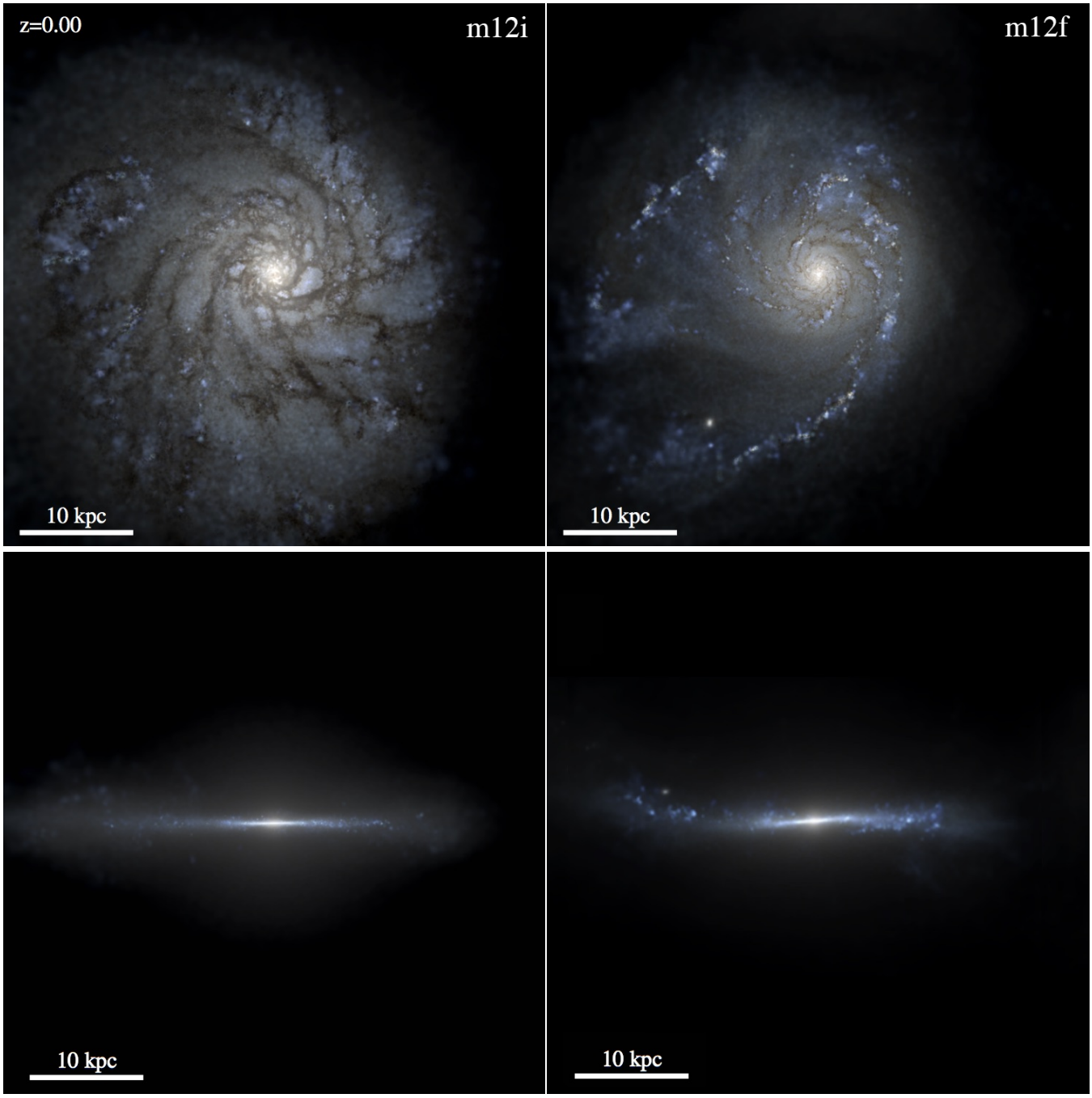

Figure 1. Mock images of two Milky Way (MW)-mass galaxies at $z = 0$ simulated using FIRE-2: (**m12i** and **m12f**). Each imaged is a $u/g/r$ composite (in Hubble Space Telescope bands) with a logarithmic stretch, using STARBURST99 to determine the SED of each star based on its age and metallicity and ray-tracing following Hopkins et al. (2005) with attenuation using a MW-like reddening curve with a dust-to-metals ratio of 0.4. We show face-on (*top*) and edge-on (*bottom*) images. Both form thin disks, with clear spiral structure, clear dust lanes, and visibly resolved star-forming regions. Properties of each galaxy (and a complete list) are in Table 1.

2009), regardless of the details of star formation in the simulation (White & Frenk 1991; Kereš et al. 2009).

However, the observed Kennicutt-Schmidt (KS) relation implies that gas consumption timescales are long (~ 50 dynamical times; Kennicutt 1998), and giant molecular clouds (GMCs) appear to turn just a few percent of their mass into stars before they are disrupted (Zuckerman & Evans 1974; Williams & McKee 1997; Evans 1999; Evans et al. 2009). Observed galaxy mass functions and the halo mass-galaxy mass relation require that galaxies incorporate or retain only a small fraction of the universal baryon fraction in stars and the ISM (Conroy et al. 2006; Behroozi et al. 2010; Moster et al. 2010). Observations of the intergalactic medium (IGM) and circum-galactic medium (CGM) require that many of those baryons must have been accreted into galaxies, enriched, and then expelled in galactic super-winds with mass loading $\dot{M}_{\rm wind}$ many times larger than the galaxy SFR (Aguirre et al. 2001; Pettini et al. 2003; Songaila 2005; Martin et al. 2010; Oppenheimer & Davé 2006), and indeed such winds are ubiquitously observed

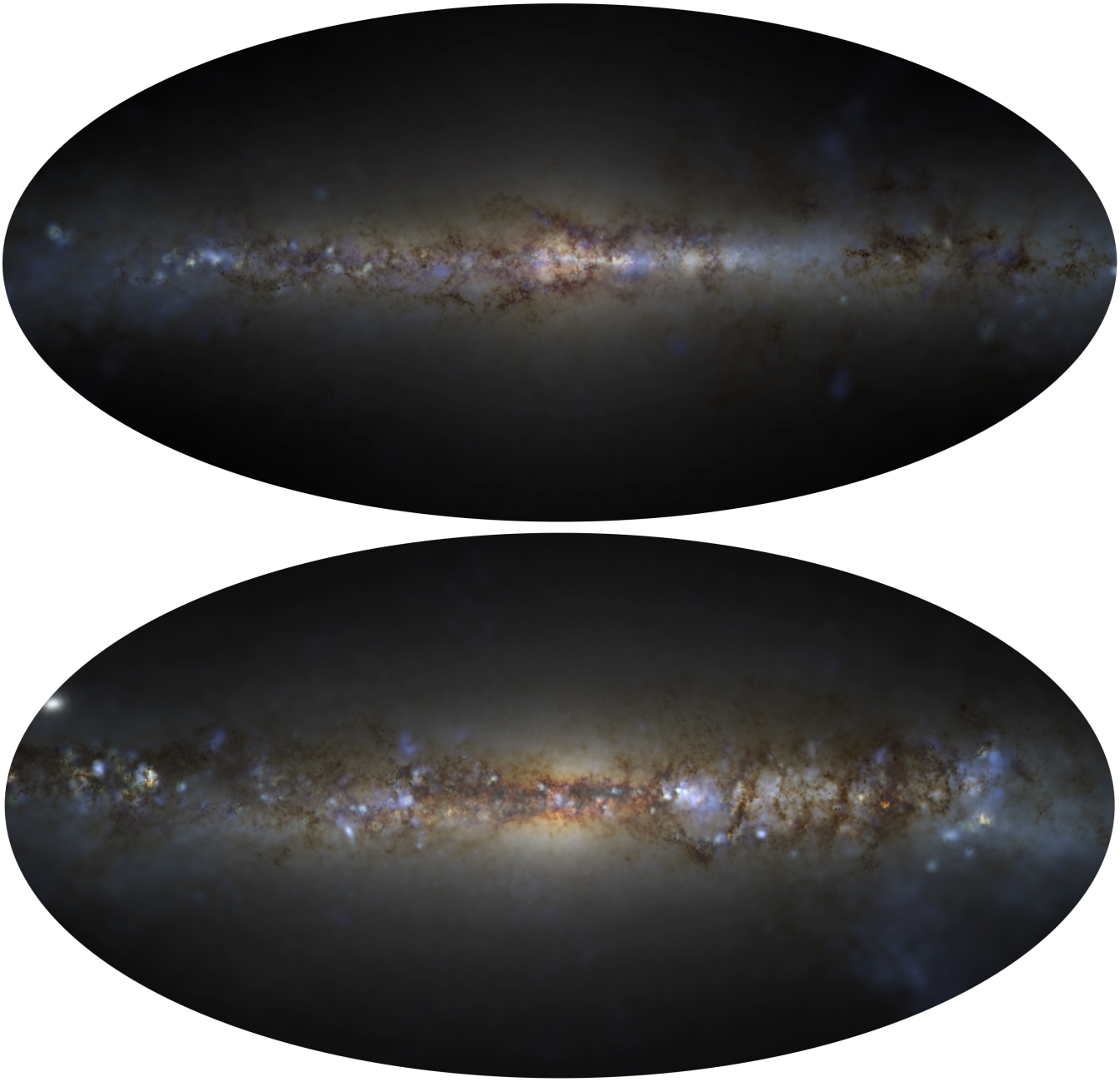

Figure 2. Mock galactic map, similar to Fig. 1, but as seen from within the galaxy, for **m12i** (*top*) and **m12f** (*bottom*). We ray-trace a Galactic (Aitoff) projection, as seen from a random star $\sim 10\,$kpc from the galactic center. Individual, filamentary giant molecular cloud (GMC) complexes and young star clusters are visible, and both galaxies have a clear thin disk plus bulge morphology.

(Martin 1999, 2006; Heckman et al. 2000; Newman et al. 2012; Sato et al. 2009; Chen et al. 2010; Steidel et al. 2010; Coil et al. 2011).

Until recently, numerical simulations treated stellar feedback in highly-simplified fashion and have had difficulty reproducing these observations. This is especially true of models which invoke only energetic feedback (thermal injection) via supernovae (SNe), which typically find the energy is efficiently radiated away (Katz 1992; Guo et al. 2010; Powell et al. 2011; Brook et al. 2011; Nagamine 2010; Bournaud et al. 2011). By "turning off cooling" for some adjusted duration, as in Stinson et al. (2006); Governato et al. (2010); Macciò et al. (2012); Teyssier et al. (2013); Stinson et al. (2013); Crain et al. (2015), or directly putting in winds "by hand" as in Springel & Hernquist (2003a); Davé et al. (2006); Anglés-Alcázar et al. (2014); Vogelsberger et al. (2014), it is possible to reproduce some of the observed galaxy properties. But this obviously does not demonstrate that known stellar feedback mechanisms actually act in this way, nor can it correctly predict many ISM and CGM-scale properties that depend explicitly on e.g. the phase-structure of feedback-driven outflows (see Hummels et al. 2013).

Accurate treatment of star formation and galactic winds ultimately requires realistic treatment of the stellar feedback processes that maintain the multi-phase ISM. Observationally, many

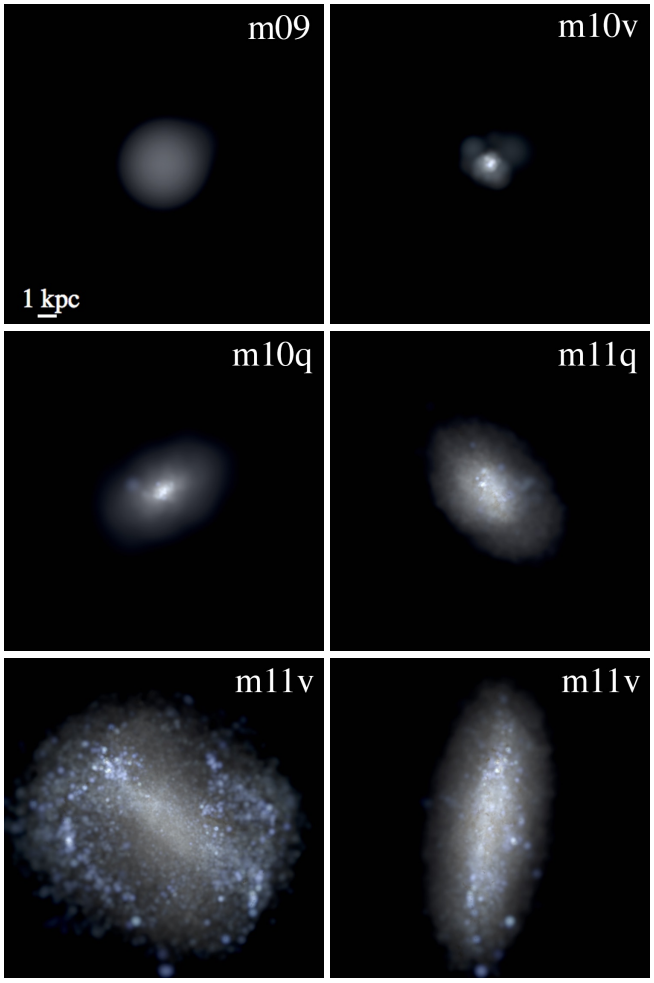

Figure 3. Mock images, as in Fig. 1, but for a subset of dwarf galaxies in our sample: **m09**, an ultra-faint with $M_* \sim 10^4\,M_\odot$ (similar to Coma Berenices, Leo IV, or Canes Venatici II); **m10v**, a faint dwarf with $M_* \sim 10^5\,M_\odot$ (similar to Hercules or Leo T); **m10q**, an intermediate-mass dwarf with $M_* \sim 10^6\,M_\odot$ (similar to Sextans, Carina, or Leo II); **m11q**, an SMC-mass dwarf with $M_* \sim 10^9\,M_\odot$; and **m11v**, an LMC-mass galaxy with $M_* \sim 2\times10^9\,M_\odot$. Most have spheroidal morphologies, as is observed and as was seen in FIRE-1 (Wheeler et al. 2015b). We show, **m11v**, the LMC-mass galaxy, both face-on and edge-on to illustrate the dramatic bar and elongated/flattened structure, similar to the actual LMC. Note that the surface-brightness scale is not the same in each image (an ultra-faint has $\sim 1000\times$ lower mean surface brightness than the LMC, so it would be invisible on the same scale).

stellar feedback processes – SNe, protostellar jets, photo-heating, stellar mass loss (O-star and AGB winds), and radiation pressure – act efficiently on the ISM (see Evans et al. 2009; Lopez et al. 2011, and references above). Simulations of either single molecular clouds/star clusters or the "first stars," which resolve individual stars and can treat these microphysics in detail, have *universally* found that the non-linear interaction of these feedback mechanisms successfully suppresses star formation, pre-process giant molecular clouds before SNe explosions (so that SNe occur in rarified environments), and generate galactic chimneys and super-bubbles that generate fountains and super-winds (e.g. Krumholz et al. 2007, 2011; Offner et al. 2009, 2011; Harper-Clark & Murray 2011; Bate 2012; Wise et al. 2012; Pawlik et al. 2013; Muratov et al. 2013). A new generation of high-resolution galaxy-scale simulations has since emerged, which reach resolution sufficient to begin directly incorporating these physics, and to begin to resolve the multi-phase structure of the ISM (Hopkins et al. 2011, 2012c; Tasker 2011; Kannan et al. 2013; Agertz et al. 2013). In a series of papers, we showed in isolated galaxy simulations that the combination of multiple feedback processes together produce a quasi-steady state ISM, in which GMCs form and disperse rapidly, with turbulence, phase structure, GMC properties, a KS law, and galactic winds in reasonable agreement with observations (see Narayanan & Hopkins 2013; Hopkins et al. 2013e, 2012a, 2013b,d,a,c).

Motivated by the success and predictive power of these simulations, in Hopkins et al. (2014) we introduced the Feedback In Realistic Environments (FIRE) project.[1] The FIRE code synthesized the physics and numerical methods developed in the previous work (with relevant improvements) into a single code suitable for high-resolution cosmological simulations of galaxy formation. These simulations explicitly treat the multi-phase ISM with heating and cooling physics from gas at a range of temperatures $T \sim 10-10^{10}\,$K, star formation restricted only to self-gravitating, self-shielding, molecular, high density ($n_{\rm H} \gtrsim 5-50\,{\rm cm^{-3}}$) gas, resolution reaching $\sim 250\,M_\odot$ or $\sim 0.5\,$pc, and (most importantly) explicit treatment of stellar feedback including the energy, momentum, mass, and metal fluxes from SNe Types I & II, stellar mass-loss (O-star and AGB), radiation pressure (UV and IR), and photo-ionization and photo-electric heating. All stellar evolution and feedback inputs are taken directly from stellar evolution models, without subsequent "parameter tuning."

In a series of papers, we have subsequently shown that cosmological zoom-in simulations incorporating these physics reproduce a diverse range of galaxy properties at a wide range of redshifts, including stellar masses, star formation histories (SFHs) and the galactic "main sequence" (Hopkins et al. 2014; Sparre et al. 2015; Feldmann et al. 2016a); metallicities and metal abundance ratios in both "standard" and r-process elements (Ma et al. 2016a; van de Voort et al. 2015); detailed morphological and kinematic structure of thin and thick disks (Ma et al. 2016d); rotation curves and morphologies of Milky Way-mass galaxies (Wetzel et al. 2016); observed satellite mass functions, rotation curves/kinematics, and cusp/core structure of dwarfs (Oñorbe et al. 2015; Chan et al. 2015; Wheeler et al. 2015a,b; Wetzel et al. 2016); abundance gradients (El-Badry et al. 2016; Ma et al. 2016b); neutral hydrogen absorption in the CGM (Faucher-Giguère et al. 2015, 2016; Hafen et al. 2016); galactic outflows (Muratov et al. 2015, 2016; Anglés-Alcázar et al. 2016); star-formation properties of galactic nuclei (Torrey et al. 2016); escape fractions of ionizing photons needed for reionization (Ma et al. 2015, 2016c); and at least some of the diversity of star-forming and quiescent massive galaxies at high redshifts (Narayanan et al. 2015; Feldmann et al. 2016a). There are of course a number of areas where the simulations fail to reproduce the observations: most notably, the bimodality of galaxy colors at both $z=0$ (Hopkins et al. 2014) and $z=2$ (Feldmann et al. 2016b) – these are likely clues to important physics missing from the FIRE-1 simulations.

For the sake of consistency and clarity, all FIRE simulations have used an *identical* source code – what we will now refer to as "FIRE-1." This ensured 100% identical physics and numerical

[1] See the FIRE project website:
http://fire.northwestern.edu
For additional movies and images of FIRE simulations, see:
http://www.tapir.caltech.edu/~phopkins/Site/animations

choices (up to the simulation resolution and choice of the specific halo simulated) in all runs, necessary for simulation comparisons. Unfortunately, this ignores development of new, more accurate hydrodynamic solvers and gravitational force softening algorithms (see e.g. Hopkins 2015), improvements to the numerical accuracy of feedback coupling algorithms (i.e. ways to ensure machine-accurate momentum conservation in SNe coupling to gas), code optimizations that would allow higher-resolution simulations, and physics neglected in FIRE-1 such as magnetic fields, cosmic rays, conduction, viscosity, optically thick radiative cooling, and more. These effects could, in principle, have large consequences for galaxy formation. For example, FIRE-1 used an improved version of the smoothed-particle hydrodynamics (SPH) method to solve the hydrodynamic equations; but it is well-known that SPH has certain low-order errors that do not converge accurately, add noise, and artificially suppress phenomena such as fluid mixing and sub-sonic turbulence (Agertz et al. 2007; Bauer & Springel 2012; Sijacki et al. 2012), potentially leading directly to large differences in cooling in massive galaxies (Kereš et al. 2012; Torrey et al. 2012). There has been considerable effort to "fix" SPH, and FIRE-1 used the improved P-SPH methods from Saitoh & Makino (2013) and Hopkins (2013a) to reduce these errors, but some (e.g. the zeroth-order errors) cannot be entirely eliminated in SPH without de-stabilizing the method (see Price 2012). As a result, especially for fluid mixing problems, newer moving-mesh or mesh-free Godunov methods provide still greater accuracy and more rapid convergence (see Springel 2010; Hopkins 2015).

We therefore introduce the "FIRE-2" simulations: an update of the FIRE physics modules in the code GIZMO. This includes a new, more accurate hydrodynamics solver that resolves the main known issues of SPH, as well as more accurate treatments of cooling and recombination rates, gravitational force softening, and numerical feedback coupling. In this paper we present a large suite of cosmological zoom-in simulations, and compare these to our FIRE-1 results and to some basic observed galaxy properties. We find that the qualitative results from the FIRE-1 simulations are reproduced in FIRE-2. We then use these simulations to extensively explore numerical and algorithmic choices in the simulation setup, and whether these have any effect on the predictions. Some first science results from these FIRE-2 simulations have been presented in Wetzel et al. (2016); Su et al. (2016); Fitts et al. (2016).

The goals of this paper are twofold. First, this is a methods and numerical/physical tests paper for the FIRE-2 simulations; we present the simulations, extensive tests of the methods, and explicitly detail all aspects of the numerical methods and algorithms. It should be possible for the reader to reproduce our methods entirely based upon this manuscript. Second, we survey numerical and physical effects, e.g.: resolution (mass, spatial, and temporal), hydrodynamic solvers (SPH vs. modern Godunov methods), criteria for star formation, details of the cooling physics, and stellar feedback from radiation, winds, and SNe. For each effect we present an extensive study in simulations to understand which effects are physical, and which numerical, and where our simulations should and should not be trusted. Because we will show that feedback is the most important property determining the galaxy's formation history, a pair of companion papers will separately explore the details of the numerical implementation and physics of mechanical/SNe feedback (Paper II; this paper is Paper I) and radiative feedback (Paper III).

Table 1 presents the initial suite of FIRE-2 simulations studied here; Table 2 provides an "executive summary" of our study and key conclusions; Table 3 provides a high-level overview of what "numerical resolution" actually means in our simulations. § 2 describes our methods: we direct the reader to the appropriate appendices where the complete algorithmic details are presented in detail. § 3 presents a basic overview of the resulting simulations and specifically examines any differences between FIRE-1 and FIRE-2 predictions. § 4 extensively studies the effects of resolution, in mass (§ 4.1), space (§ 4.2), and time (§ 4.3). § 5 examines the effects of the hydrodynamic methods, including SPH vs. finite-volume methods and § 6 considers the effects of so-called "artificial pressure" terms used in some (non-FIRE) simulations. § 7 studies the details of cooling, and § 8 considers the star formation algorithm. § 9 considers the effects of different stellar feedback physics, turned on and off in turn, to provide an indication of which feedback processes dominate, and to provide a way of quantifying the relative importance of numerical and physical (feedback) uncertainties for our results. The Appendices present additional cooling/feedback tables and algorithmic information necessary for implementing the FIRE-2 simulations.

2 METHODS

Here we describe our numerical implementation of hydrodynamics, gravity, cooling, star formation, stellar feedback, and additional fluid physics in the FIRE-2 simulations. For further details, at the end of each sub-section below, we direct interested readers to the appendix, paper, or public code where an exact algorithmic breakdown of our numerical methods is provided: these should be sufficient to completely reproduce the results here. We will study, test, and compare each aspect of the methodology in more detail below.

Before FIRE-1, a series of papers developed the numerical methods, and tested each physical addition individually using much higher-resolution ISM-scale simulations and exact analytic solutions where possible (we refer the interested reader to Hopkins et al. 2011, 2012c,b,a, 2013a,b,c,e,d). These developments and improvements were then synthesized into the physics implemented in FIRE-1, described in detail in Hopkins et al. (2014).

In our FIRE-2 runs, the "core" or "baseline" physics is the same as in FIRE-1: we simply seek to improve the numerical accuracy with which we solve the relevant equations. However, the switch to a new hydrodynamics method in FIRE-2 also makes it possible *in principle* to include new physics such as magnetic fields: these are described here as "additional" physics, and will be studied in separate work.

As with FIRE-1, all runs denoted as "FIRE-2" here or in any other papers (e.g. Wetzel et al. 2016; Su et al. 2016; Fitts et al. 2016) use the *identical* source code and physical parameters, unless explicitly labeled otherwise for comparison purposes. Of course certain numerical parameters (e.g. force softening) scale explicitly with resolution; these are provided for each simulation.

2.1 Hydrodynamics

A major motivation for our introduction of "FIRE-2" is to take advantage of a new generation of accurate, mesh-free Godunov hydrodynamics methods that have been recently developed (see Gaburov & Nitadori 2011; Hopkins 2015). Because we enforced the strict requirement that all FIRE-1 simulations use the identical source code, all FIRE-1 runs used the older "P-SPH" method (Hopkins 2013a), an improved "pressure-energy" variant of smoothed-particle hydrodynamics (SPH). Below, we will explore the effects this has on the pure hydrodynamics in our simulations. More importantly, however, the new Godunov methods allow us to accurately include more complicated plasma physics such as magnetic

fields and anisotropic diffusion, which were not possible to solve accurately in P-SPH (see Hopkins & Raives 2016; Hopkins 2016b).

We therefore employ the meshless finite-mass (MFM) magneto-hydrodynamics solver in GIZMO.[2] This is a mesh-free, Lagrangian finite-volume Godunov method designed to capture advantages of both grid-based and particle-based methods, built on the gravity solver and domain decomposition algorithms of GADGET-3 (Springel 2005). In a series of methods papers (Hopkins 2015; Hopkins & Raives 2016; Hopkins 2016a,b), GIZMO has been tested extensively (involving ~ 100 distinct test problems) compared to state-of-the-art fixed grid Godunov codes (e.g. ATHENA and RAMSES; Teyssier 2002; Stone et al. 2008), moving-mesh codes (e.g. AREPO; Springel 2010), and "modern" SPH methods (e.g. P-SPH; Hopkins 2013a; Rosswog 2014; Hu et al. 2014).

We emphasize that in essentially every test problem we find MFM gives more accurate results and faster convergence compared to state-of-the-art SPH methods, and demonstrates accuracy and convergence in good agreement with well-studied fixed-grid and moving-mesh codes. Most importantly, this includes areas where SPH has had historical difficulty, including sharp shock-capturing, fluid-mixing instabilities, magneto-rotational instabilities, and anisotropic diffusion (Ritchie & Thomas 2001; Agertz et al. 2007; Price 2008; Wadsley et al. 2008; Read & Hayfield 2012; Saitoh & Makino 2013). For some problems relevant in cosmological simulations, e.g. those with moving contact discontinuities, orbiting thin disks, supersonically shearing fluid-mixing instabilities, poorly resolved explosions, hydrostatic gravitational equilibrium or gravitational collapse, the Lagrangian nature of the method here also allows us to converge at much lower resolution compared to fixed-grid methods (Müller & Steinmetz 1995; Zingale et al. 2002; O'Shea et al. 2005; Heitmann et al. 2008; Hopkins 2015) and provides excellent angular momentum conservation (avoiding "grid alignment" and spurious torques common in grid-based codes; Hahn et al. 2010; Byerly et al. 2014; Hopkins 2015).

As discussed in Hopkins (2015), this increased accuracy and convergence rate effectively makes our simulations effectively higher-resolution (at least in terms of the spatial resolution of the hydrodynamics and its convergence), compared to FIRE-1 simulations at the same particle number.

For reasons discussed in § 6, we do not adopt an artificial "pressure floor" of any kind for hydrodynamics; unresolved fragmentation is instead explicitly treated via our star formation model.

Readers interested in further details of the hydrodynamic solver should consult Hopkins (2015) and the public GIZMO source code.

2.2 Gravity

The N-body gravity solver is extensively detailed in Hopkins (2015); this is an improved version of the GADGET-3 Tree-PM solver and additional details can be found in Springel (2005). This solver is well-tested under a huge range of applications; we will focus here only on how this relates to the spatial or force resolution, and what this means for "resolved scales" in FIRE.

By default, as described in Hopkins (2015), the resolution of gravity and hydrodynamics are equal and the two use the *same*, consistent definition and assumptions about the gas mass distribution in the simulations. Specifically, we follow Price & Monaghan (2007) and compute gravitational forces from gas particles by assuming the gas is in an extended mass distribution with the same functional form as the interaction kernel centered at the particle; Hopkins (2015) show this is the leading-order accurate expression for the potential if we integrate Poisson's equation using the exact differential mass distribution assumed in the hydrodynamic equations (it is second-order accurate averaged in shells around the particle center). This means that, for gas, the gravitational force softening resolution $\epsilon_i^{\rm gas}$ identically follows the inter-particle/cell separation $h_i^{\rm gas} = \Delta x$, where Δx is the equivalent cell-length in a fixed, Cartesian mesh (i.e. the average distance between particle centers, around element i, is $h_i^{\rm gas}$), so $\epsilon_i^{\rm gas} \equiv h_i^{\rm gas} = 16\,{\rm pc}\,m_{i,\,1000}^{1/3}\,(n_{\rm H}/10\,{\rm cm^{-3}})^{-1/3}$ (where $m_{i,\,1000}$ is the mass resolution in units of $1000\,M_\odot$). Note this definition is *independent* of the exact softening kernel shape or "neighbor number." Exact definitions of "force resolution," and values for $h_i^{\rm gas}$, are discussed extensively in § 4.2 below. Because the separations $h_i^{\rm gas}$ change, we must be careful to maintain energy and momentum conservation correctly when elements interact inside the softening; Price & Monaghan (2007) show how the relevant expressions can be rigorously derived from the softened N-body particle Lagrangian and the expressions used in GIZMO are presented in Hopkins (2015), Appendix H2.

A number of detailed studies have shown this produces the "optimal" softening in terms of (i) physical consistency with the hydrodynamics, (ii) maximizing accuracy and convergence rates, and (iii) reducing N-body integration and force errors (e.g. Merritt 1996; Bate & Burkert 1997; Romeo 1998; Athanassoula et al. 2000; Dehnen 2001; Rodionov & Sotnikova 2005; Price & Monaghan 2007; Barnes 2012; Hubber et al. 2013). Some simulations – unlike FIRE – adopt a much larger gravitational force softening compared to the hydrodynamic resolution. Such a choice is always unphysical, because the Poisson equation is being solved for a different mass distribution, compared to the one the hydrodynamic equations are solving (Barnes 2012). In addition to larger integration errors and an artificial loss of resolution, the mis-match in resolution can produce unphysical outcomes that do not converge correctly (see e.g. Bate & Burkert 1997). Explicit tests validating the accuracy and convergence of our adaptive self-gravity implementation in GIZMO are presented in Hopkins (2015), including the Evrard (1988) polytropic collapse test, the gas (and gas+DM) Zel'dovich (1970) pancake collapse, the "Santa Barbara Cluster" adiabatic zoom-in simulation (Frenk et al. 1999), and rotating steady-state stable disk tests.

For dark matter, the collisionless nature of the fluid makes the "correct" softening ambiguous. Unlike a gas cell, which cannot "move through" a neighbor gas cell, and therefore must spread its mass over a larger/smaller volume when $\nabla\cdot\mathbf{v} \neq 0$, collisionless fluids do not obey this relation and therefore do not necessarily obey any unique relation between the real-space DM N-body particle distribution and implied phase-space distribution of DM (for discussion, see Abel et al. 2012). We therefore have two choices which we explore in detail below. In our default simulations, we simply set $\epsilon_i^{\rm DM} = $ constant. We will consider extensive tests of the exact DM softening, and show it has almost no effect on our conclusions, but still choose values that maximize integration accuracy and convergence. Alternatively, we adopt adaptive softenings for the dark matter, using the same definition of the neighbor kernel as for gas. In other words, we set $\epsilon_i^{\rm DM} = h_i^{\rm DM}$, where $h_i^{\rm DM}$ is determined from the *dark matter* particle neighbor distribution in the identical manner to $h_i^{\rm gas}$, ensuring $h_i^{\rm DM}$ is the mean (kernel-averaged) inter-particle spacing *of dark matter particles* within the kernel.

[2] A public version of this code is available at http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html

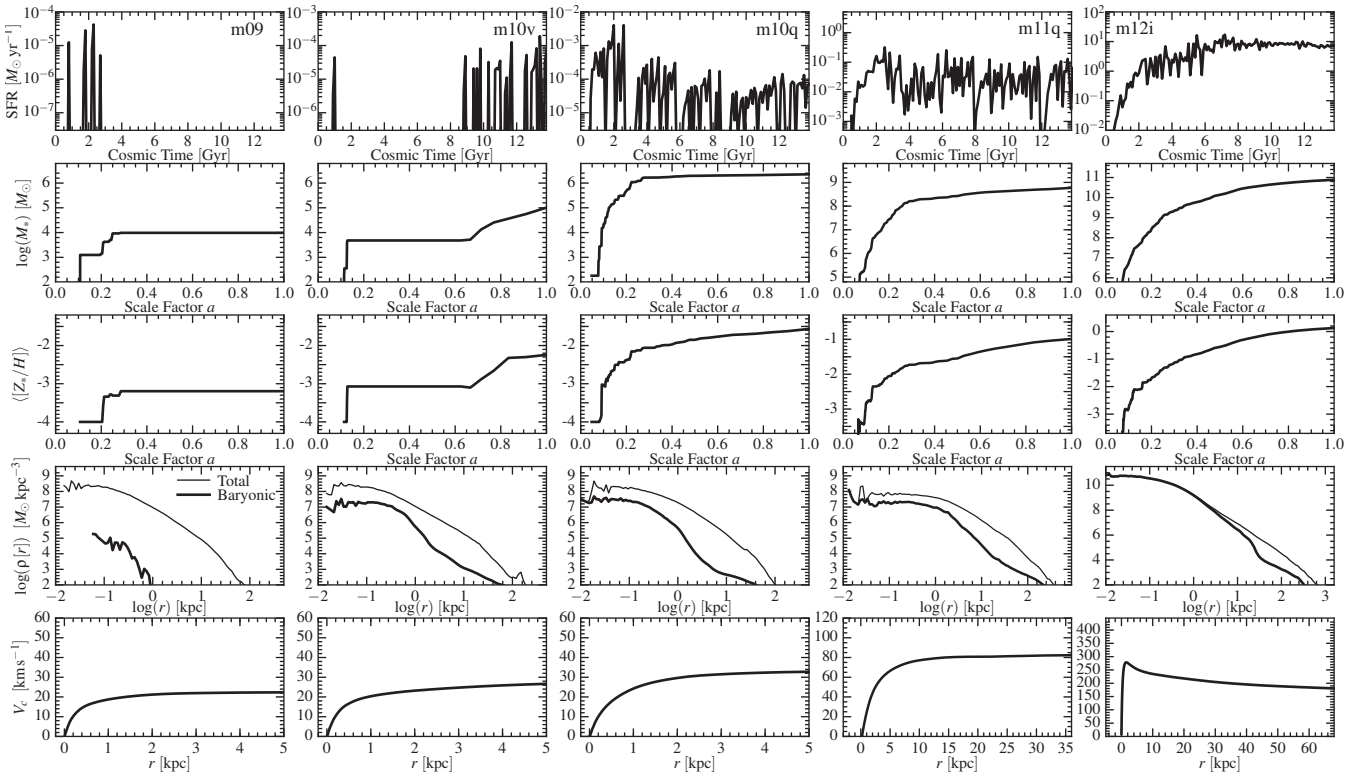

Figure 4. Basic galaxy-scale properties in a subset of FIRE-2 simulations. Each column shows a different galaxy from Fig. 1 and Table 1. *Top:* Star formation rate (averaged in 100 Myr intervals) of the primary (most massive) galaxy in each simulation versus cosmic time. Ultra-faint galaxies ($V_c \lesssim 20\,\mathrm{km\,s^{-1}}$) are quenched after reionization. In more massive dwarf galaxies, SF is highly "bursty", but it becomes less so at even higher masses and at later times (as seen in our FIRE-1 runs; Sparre et al. 2015). *Second from Top:* Total stellar mass versus scale factor, $a = 1/(1+z)$, in the zoom-in region (we show against scale factor so that the rate of growth at early times is more clear). We show all stellar mass within $R_{\rm vir}$, but this is dominated by the main galaxy, so it evolves similarly. Growth occurs rapidly at high redshifts then settles into a more steady state at late times, allowing galaxy structure to relax (El-Badry et al. 2016). *Middle:* Stellar Mass-Weighted Average Metallicity versus scale factor. This rise is similar to the stellar mass, because these galaxies evolve on a redshift-dependent stellar mass-metallicity relation; the metallicities at each mass and redshift are nearly identical to FIRE-1 galaxies (see Ma et al. 2016a). *Second from Bottom:* Baryonic (*thick*) and total (*thin*) mass density profiles as a function of radius around the most massive galaxy at $z = 0$. Profiles are averaged in spherical shells. The dwarfs with stellar masses $M_* \sim 10^6 - 10^{9.5}\,M_\odot$ exhibit central "cores" in their mass profiles, in both stars and dark matter, most prominent at $M_* \sim 10^9\,M_\odot$, where the cores extend to $\sim$ kpc scales, consistent with FIRE-1 results (Oñorbe et al. 2015; Wheeler et al. 2015a; Chan et al. 2015). *Bottom:* Rotation curves of circular velocity, V_c, versus radius around the most massive galaxy at $z = 0$. Dwarfs exhibit slowly-rising rotation curves, while MW-mass systems have flat rotation curves with small-to-modest bulges, as in FIRE-1 (Chan et al. 2015).

This gives $\epsilon_i^{\rm DM} \equiv h_i^{\rm DM} = 40\,\mathrm{pc}\ m_{i,\,1000}^{1/3}\,(\rho_{\rm DM}/10^8\,M_\odot\,\mathrm{kpc^{-3}})^{-1/3}$. We will show the choice of adaptive or fixed DM softenings has little effect, consistent with previous studies (Bagla & Khandai 2009; Iannuzzi & Dolag 2011; Iannuzzi & Athanassoula 2013). However, we will show that adaptive DM softenings require more stringent timestep criteria for the DM, which can make them computationally inefficient.

For stars, the "correct" softening is even more ambiguous, since they act as a collisionless fluid in some circumstances, but in others are collisional, and do cluster strongly. We therefore also default to a simple constant softening, matched to the gas softening at the mean gas density of star formation. But we also compare adaptive-softening runs, where we set h_i^* using the distribution of all baryonic particles around the star (i.e. gas and stars, not just other stars).[3] The choice of stellar softening does not appear to affect our results.

Readers interested in further details of the gravity solver and adaptive softening scheme should consult Hopkins (2015) and the public GIZMO source code.

2.3 Cooling

Gas cooling is solved using a standard implicit algorithm, described in Hopkins et al. (2014). Heating/cooling rates are computed including free-free, photo-ionization/recombination, Compton, photo-electric, metal-line, molecular, fine-structure, dust colli-

[3] The use of all baryons to determine h_i^* avoids two clearly un-physical limits. If we used only other stars, then the first stars forming in a galaxy would dis-continuously "jump" from small $h_i^{\rm gas}$ (from the high-density gas in which they formed) to arbitrarily large h_i^* encompassing the whole galaxy! However, if we used only gas to determine h_i^*, then old stars in a gas-depleted disk might be artificially softened over a huge radius encompassing many additional stars. Instead, our choice guarantees h_i^* evolves reasonably continuously from the progenitor gas $h_i^{\rm gas}$ at the time of star formation, through to the inter-star neighbor distance in gas-poor stellar systems.

sional, and cosmic ray processes, from $T = 10 - 10^{10}\,$K. We follow 11 separately-tracked species (H, He, C, N, O, Ne, Mg, Si, S, Ca, and Fe, each with its own yield tables associated directly with the different mass return mechanisms below; see Wiersma et al. 2009b). The relevant ionization states are tabulated from CLOUDY simulations including the effects of uniform but redshift-dependent background (from Faucher-Giguère et al. 2009) together with local radiation sources (every star particle is treated as a source; see the feedback description below). We account for self-shielding with a local Sobolev/Jeans-length approximation (integrating the local density at a given particle out to a Jeans length to determine a surface density Σ, then attenuating the flux seen at that point by $\exp(-\kappa_\nu\,\Sigma)$); this approximation has been calibrated in full radiative transfer experiments in Faucher-Giguère et al. (2010) and Rahmati et al. (2013).

With this, high-temperature ($> 10^4\,K$) metal-line excitation, ionization, and recombination rates then follow Wiersma et al. (2009a). Free-free, bound-free, and bound-bound collisional and radiative rates for H and He follow Katz et al. (1996) with the updated fitting functions in Verner & Ferland (1996). Photo-electric rates follow Wolfire et al. (2003), accounting for PAHs and local variations in the dust abundance. Compton heating/cooling (off the combination of the CMB and local sources) follows Faucher-Giguère & Quataert (2012). Fine-structure and molecular cooling at low temperatures ($5 - 10^4\,$K) follows a pre-computed tabulation of CLOUDY runs as a function of density, temperature, metallicity, and local radiation background (see Robertson & Kravtsov 2008). Collisional dust heating/cooling rates follow Goldsmith & Langer (1978) with updated coefficients from Meijerink & Spaans (2005) assuming a minimum grain size of 10 Å, and a dust temperature of 30 K. Cosmic ray heating follows Guo & Oh (2008) accounting for both hadronic and Compton interactions, with a uniform cosmic ray background of $\sim 5\,\mathrm{eV\,cm^{-3}}$. At very high densities ($\sim 10^{10}\,\mathrm{cm^{-3}}$), gas can become optically thick to its own cooling radiation; this is treated self-consistently following Rafikov (2007), but this is irrelevant for the simulations here (because they do not reach sufficiently high densities). Hydrodynamic heating/cooling rates (including shocks, adiabatic work, reconnection, resistivity, etc.) are computed in standard fashion by the hydro solver; this is incorporated directly into our fully-implicit solution rather than being operator-split, since operator-splitting can lead to large errors in temperature in the limit where the cooling time is much faster than the dynamical time. A 10 K temperature floor is enforced, but this has no detectable effect on our conclusions.

For readers interested in reproducing our results, we provide simple fitting functions to the complete set of cooling physics above, in Appendix B.

2.4 Star Formation

Gas which is locally self-gravitating, self-shielding, Jeans unstable, and above some minimum density is turned into stars using a sink-particle approach. Specifically, gas is eligible to turn into stars if and only if it meets the following criteria:

(i) **Self-Gravitating:** We require the potential energy be larger than the thermal plus kinetic energy within the resolution scale: specifically we use the sink-particle criterion developed in Hopkins et al. (2013d), $\alpha \equiv (\delta v^2 + c_s^2)\,\delta r / G m_{\rm gas}(<\delta r) = [\|\nabla\otimes\mathbf{v}\|_i^2 + (c_{s,i}/h_i)^2]/(8\pi G\rho_i) < 1$, where $\delta v = \|\nabla\otimes\mathbf{v}\|_i\,h_i$ and c_s give the kinetic and thermal energy, respectively, within the resolution scale $\delta r \rightarrow h_i$ around the particle ($\otimes$ is the outer product). Hopkins et al. (2013d) and many other studies (Li et al. 2005; Federrath et al. 2010; Padoan & Nordlund 2011; Padoan et al. 2012) have shown this is more useful than a density criterion (as well as more accurate in converging to the results of higher-resolution simulations), since it actually identifies gas which is collapsing under self-gravity at the resolution scale (i.e. the gas which should, physically form stars), *independent* of the exact spatial, mass, or density scale. This also does not allow unbound material to form stars (e.g. tidally unbound gas, or gas in strong shocks and winds which is dense, but not self-gravitating owing to large internal motions). The exact order-unity coefficient is calibrated from higher-resolution simulations of collapsing clouds/cores (see Padoan et al. 2012; Federrath & Klessen 2012), but our results are insensitive to variations.

(ii) **Self-Shielding (Molecular):** We estimate the self-shielded ("molecular") fraction of each gas element following Krumholz & Gnedin (2011), using the local Sobolev approximation and metallicity to estimate the integrated column to dissociating radiation, and allow star formation only from the molecular component ($\rho_{\rm mol} = f_{\rm mol}\,\rho$). This is specifically a requirement that the gas be self-shielding, and therefore able to cool to low temperatures. Given the high $n_{\rm crit}$ (see below), this criterion typically has negligible impact, since the high-density gas is typically all shielded and molecular anyways.

(iii) **Jeans Unstable:** We require a thermal Jeans mass below the maximum of the particle mass or $10^3\,M_\odot$ in the element. This is done to ensure that any resolved, massive self-gravitating objects which should collapse coherently (as opposed to fragmenting into stellar mass-scale objects) are followed self-consistently and not simply assigned to stars (the choice of $10^3\,M_\odot$ is designed to separate massive "clumps" from normal very massive stars, but is not important). In practice this criterion is always easily met when other criteria are met.

(iv) **Dense:** To prevent spurious application of the criteria above, we also check that $n_{\rm H} > n_{\rm crit} = 1000\,\mathrm{cm^{-3}}$ (much larger than the mean galaxy density $\langle n\rangle$). This restricts star formation to dense molecular clouds fragmenting out of the background disk.

If gas meets all the criteria above, we assume it turns into stars at a rate $\dot{\rho}_* = \rho_{\rm mol}/t_{\rm ff}$ where $t_{\rm ff}$ is its free-fall time. This also comes directly from higher-resolution simulations of turbulent clouds (Padoan & Nordlund 2011; Padoan et al. 2012), as well as analytic models for star formation via turbulent fragmentation (Hopkins 2012b,a; Guszejnov & Hopkins 2015, 2016; Guszejnov et al. 2016). We stress that this is an assumption about the rate at which small, *locally self-gravitating* clumps (which may only be a small fraction of the dense gas mass) fragment; it does *not* imply the global efficiency of star formation (either on galaxy or GMC scales) is necessarily high – we find that it is self-regulated by feedback at $\sim 1-10\%$ per free-fall time (see Hopkins et al. 2011, 2013e, 2014) even in gas with densities $\gtrsim 100\,\mathrm{cm^{-3}}$, in excellent agreement with observations. Recently, similar implementations to ours have also found consistent results on large scales (Semenov et al. 2015; Agertz & Kravtsov 2016), consistent with analytic expectations (Ostriker & Shetty 2011; Faucher-Giguère et al. 2013). If particles do not meet all of the criteria above, their SFR is zero. Gas particles which turn into star particles begin life as zero-age main sequence populations, with abundances and total mass inherited from their progenitor gas particle.

We provide the complete set of formulae and detailed algorithmic implementation of star formation in Appendix C.

2.5 Stellar Feedback

Once a star particle forms, it is treated as a single stellar population, with known age $t_* = t - t_{\rm form}$, metallicity (inherited from its progenitor gas particle), and mass (equal to its progenitor gas particle). All feedback quantities are tabulated directly – *without subsequent adjustment or fine-tuning* – from standard stellar population models (STARBURST99; Leitherer et al. 1999) assuming a Kroupa (2001) IMF (the same as FIRE-1).[4]

Here, we briefly summarize the feedback mechanisms. Because these are the most important and novel aspect of the FIRE simulations, we discuss the exact physics and algorithmic implementation in much greater detail in the companion papers, Paper II & Paper III. These papers present extensive tests of the algorithms, with idealized simulations of e.g. individual SNe explosions reaching resolution $< 0.01\,M_\odot$ and experiments using full radiation-hydrodynamics simulations, used to validate the exact implementations here. But for the sake of completeness, we summarize them here and provide the complete algorithms in the Appendices.

The physics of stellar feedback in FIRE-2 are the same as in FIRE-1 (Hopkins et al. 2014), and the algorithms are identical up to improvements in accuracy which we explicitly detail below.

(i) **Supernovae (Ia & II):** Every timestep Δt, for each star particle, the tabulated SNe rate as a function of star particle mass, age, and metallicity is used to determine the probability p of an event (Type-Ia and/or Type-II) occurring within the particle within Δt; our mass and time resolution is such that $p \ll 1$, i.e. we explicitly treat individual SNe explosions, rather than model their collective effects indirectly. We determine probabilistically if an event occurs within Δt; if so, the appropriate ejecta mass, metal yields, energy, and momentum (also determined from the stellar evolution tables) are deposited directly in the surrounding gas around the star particle. The algorithm for deposition is constructed to ensure *manifest, machine-accurate* conservation of mass, metal mass, energy, and momentum, while also ensuring that the ejecta are distributed isotropically in the rest frame of the star. In Paper II we show that this is non-trivial in Lagrangian codes such as ours, where highly anisotropic gas distributions around a star particle can easily bias the momentum distribution and even violate linear momentum conservation, if the algorithm is not carefully designed to prevent this. We properly account for the relative star-gas motion (so e.g. the exact shock solution includes the initial stellar motion through the background gas). We determine the coupled momentum by computing the exact Sedov-Taylor solution for an energy-conserving spherical shock, at the coupling location (resolved separation between gas and star); if the resulting momentum exceeds the terminal momentum at which point the shock should have become radiative (equivalently, if the resolved coupling radius is larger than the cooling radius), we deposit only the momentum which would have been present when it reached that cooling radius, and assume the remaining thermal energy has been radiated away. Paper II shows that this ensures our simulations exactly reproduce the fully-converged solutions (with resolution $< 0.01\,M_\odot$) for individual SNe explosions in high-resolution ISM simulations (once they reach the same radius as our coupling radius), independent of our resolution, for the same ambient density. We stress that we do not turn off cooling or otherwise impose any assumption about "galactic wind driving."

(ii) **Continuous Stellar Mass-Loss (OB/AGB-star Winds):** Similarly, stellar mass-loss is injected continuously in the gas surrounding each star particle as a function of stellar age and metallicity, with the appropriate mechanical luminosity, momentum, mass, and metal content, including both fast (O/B-star) and slow (AGB) winds, calculated directly from the stellar evolution models. The algorithm for deposition is exactly the same as for SNe, except there is an "event" every timestep with associated ejecta mass $= \Delta t\,\dot{M}_{\rm wind}$.

(iii) **Photo-Ionization and Photo-Electric Heating:** For computing radiative feedback properties, each star particle is treated as a source with an appropriate age and metallicity-dependent, IMF-averaged spectrum. We approximate the complete spectrum with a five-band treatment that includes ionizing, far-UV (relevant for photo-electric heating), near-UV, optical/near-IR, and mid/far-IR (re-radiated dust emission) photons. The background radiation owing to these sources is then locally extincted by the gas immediately surrounding the star (using a Sobolev approximation to estimate the column integrated to infinity, and extincting each band accordingly), with frequency and metallicity-dependent opacities from dust and neutral gas. The luminosity absorbed by dust (non-ionizing bands) is assumed to re-radiate in the mid/far-IR band. The resulting, post-extinction luminosities are then propagated to long-range distances through an optically thin transport network (using a tree structure), to calculate an incident flux in each band at all positions. We therefore refer to this transport algorithm as the "Locally Extincted Background Radiation in Optically-thin Networks," or LEBRON, approximation. Since we simulate only a single small region of the Universe surrounding one galaxy in our "zoom-in" simulations, we add to the diffuse ionizing-band flux a uniform but redshift-dependent meta-galactic background tabulated from Faucher-Giguère et al. (2009). The fluxes are then corrected for self-shielding using the same local-extinction Sobolev approximation, at the location of the gas. The resulting incident ionizing and FUV fluxes are then used to self-consistently compute the gas ionization states and radiative heating/cooling rates in our standard cooling algorithms described in Appendix B.

(iv) **Radiation Pressure:** As photons are tracked according to the algorithm above, each *explicitly resolved* absorption transfers the appropriate photon momentum $= L_{\rm abs}\,\hat{n}/c$ (where $\hat{n}$ is the direction of ray propagation) to the gas. This automatically accounts for both direct UV/optical single-scattering, and indirect re-radiated IR photons (which can, in principle, be multiply-scattered, although this rarely occurs on the relatively coarse scales we resolve in the FIRE simulations). We stress that we *do not* assume any "sub-grid" photon coupling, multiple-scattering, or radiation pressure – there is no "boost factor" anywhere in the model, nor has there ever been one: in both FIRE-2 and FIRE-1, the only radiation pressure in the simulations is from explicitly resolved photon absorption. In Paper III, we show that only $\sim 1/2$ of the total bolometric luminosity of stellar populations is absorbed at all, and, given our numerical resolution (which prevents us from resolving e.g. proto-stellar cores), the multiple-scattering IR term accounts for $< 10\%$ of the galaxy-averaged radiation pressure (it may be important, however, in dense galactic nuclei corresponding to observed systems like Arp 220).

We emphasize that while quantities like SNe rates and stellar spectra are IMF-averaged, individual SNe are always discrete events (not continuous energy injection). In future work we will consider the effects of explicitly sampling the spectrum of stellar

[4] Of course, alternative stellar evolution/IMF models may predict different feedback properties, but we will not investigate this here. In general the predicted variation is small, but for some quantities, e.g. the escape fraction of ionizing photons at high redshift, we have shown it can be important (Ma et al. 2016c).

masses from the IMF (Su et al., in prep.); however our preliminary results indicate the effects on large scales are (unsurprisingly) small compared to our IMF-averaged approach.

For readers interested in reproducing our results, we provide simple fitting functions to all of our stellar evolution tabulations (and yield tables) needed for the feedback mechanisms above, in Appendix A. All details of the algorithmic implementation of mechanical feedback (SNe and stellar mass-loss) are given in Appendix D, and all details of the algorithmic implementation of radiative feedback (radiation pressure, photo-ionization, and photo-electric heating) are given in Appendix E.

2.6 "Additional" Physics: Magnetic Fields, Conduction, Viscosity, Diffusion, Cosmic Rays, Black Holes, and more

As noted above, a major motivation for our migration to FIRE-2, using the new MFM hydrodynamic solver, is to compare simulations including more complicated plasma physics, e.g. magnetic fields. As shown in Hopkins & Raives (2016) and Hopkins (2016b), SPH (at least as implemented in GIZMO) has fundamental low-order errors which converge slowly and, in some cases, corrupt the correct solutions for certain anisotropic diffusion and magnetic effects even at arbitrarily high resolution.

However, for the sake of clarity and direct comparison with FIRE-1, in this paper we will focus on simulations that include our "core" set of FIRE physics (gravity, hydrodynamics, cooling, star formation, and stellar feedback, as described above). This means that our "default" or "core physics only" FIRE-2 simulations use *the same physics* as FIRE-1, just more accurate numerical integration of those physics. The effects of additional physics will of course be the subject of their own studies.

Some examples of these "additional physics" include:

(i) Magnetic fields (§ F1): The equations of ideal magnetohydrodynamics (MHD) are implemented and tested for our MFM solver in Hopkins & Raives (2016) (see also Hopkins 2016a). Non-ideal terms (Ohmic resistivity, ambipolar diffusion, and the Hall effect) are also implemented and tested in Hopkins (2016b). A preliminary exploration of the effects of magnetic fields on galaxy properties is presented in Su et al. (2016).

(ii) Anisotropic (Braginskii) Conduction & Viscosity (§ F2): The fully-anisotropic equations of Spitzer-Braginskii conduction and viscosity are implemented and tested in Hopkins (2016b). A first exploration of these physics in FIRE-2 simulations is presented in Su et al. (2016).

(iii) Passive-Scalar Turbulent Eddy Diffusivity (Metal Diffusion) (§ F3): Because our default hydrodynamic method follows finite-mass elements, micro-physical diffusion of passive scalars (e.g. metals) in a turbulent ISM requires an explicit numerical diffusion term. This is implemented following the standard diffusion operators in Hopkins (2016b) and described in Appendix F. Tests of its effects in FIRE-2 simulations are shown in § 7.2; it will be explored in more detail in Escala et al. (in prep.).

(iv) Cosmic Rays (§ F4): Cosmic ray physics including injection by SNe, fast stellar winds and shocks; hadronic and Compton losses; fully-anisotropic streaming, diffusive, advective and adiabatic transport and heating coupled to magnetic fields, are implemented in Hopkins (2016b) and Appendix F. A first exploration of their effects in FIRE-2 simulations will be presented in Chan et al. (in prep.).

(v) Alternative Radiation Hydrodynamics (§ F5): All FIRE-2 simulations include some approximate radiation-hydrodynamics treatment. Our default method uses the LEBRON algorithm described in § 2.5; however we have also implemented and tested a variety of additional radiation hydrodynamics treatments including the FLD, OTVET, and M1 moments methods and explicit ray-tracing methods. These are presented and compared in more detail in Paper III and Khatami et al. (in prep.).

(vi) Super-Massive Black Holes (§ F6): FIRE-2 includes (optional) numerical modules for the treatment of supermassive black holes (BHs) as sink particles, including their seeding; growth via gas accretion, capture of collisionless particles, or BH-BH mergers; and feedback in the form of radiation (radiation pressure, Compton and photo-ionization heating), mechanical work (accretion disk winds and jets), and cosmic ray injection. Each aspect of these physics will be explored separately in future work (e.g., Anglés-Alcázar et al., in prep.).

Details of the numerical implementation of the physics above are given in Appendix F. Interested readers should also consult the relevant methods papers and the public GIZMO source code.

2.7 Timesteps & Integration

Our time integration scheme is discussed in detail in Hopkins (2015). Following Springel (2005) we use an adaptive power-of-two hierarchy for assigning individual timesteps for particles. As shown in Saitoh & Makino (2009) and Durier & Dalla Vecchia (2012), in problems with high Mach number flows, adaptive timesteps can lead to errors if particles with long timesteps interact suddenly mid-timestep with those on much shorter timesteps; this is easily remedied by requiring that, at all times, any active particle informs its neighbors of its timesteps and none are allowed to have a timestep > 4 times that of a neighbor. Whenever a timestep is shortened (or energy is injected in feedback of any sort) particles are forced to return to the timestep calculation. This has been tested extensively in Hopkins (2013a) and Hopkins (2015).

The timestep is set by the minimum of various criteria. All particles obey limits $\Delta t < 0.2\,(h_i/|\mathbf{a}_i|)^{1/2}$ (see Power et al. 2003; here h_i is the minimum of the inter-particle separation or Plummer-equivalent force softening)[5] and $\Delta t < 0.25/|\nabla\cdot\mathbf{v}_i|$, where $\mathbf{v}_i$ and $\mathbf{a}_i$ are the total velocity and acceleration of particle i (including all sources of acceleration: e.g. feedback and hydrodynamic forces, for gas). For further safety, we always enforce a maximum timestep of $\Delta a < 10^{-4}\,a$ (where a is the scale-factor), but this is almost never important. Gas elements must also obey the Courant (CFL) condition: $\Delta t < 0.4\,h_i/v_{\rm sig,\,i}^{\rm max}$, where $v_{\rm sig,\,i}^{\rm max}$ is the usual maximum signal velocity between all particles interacting with i (see Hopkins 2015 for tests and details). Additional timestep criteria apply if additional fluid physics (magnetic fields, diffusion, cosmic rays, radiation) are included (see Hopkins 2016b).[6] In the above equations, note h_i is defined by the inter-particle separation as defined in § 4.2, which is the appropriate value for the prefactors here (see Hopkins 2015; the pre-factors would need to decrease by a factor ~ 2 if we replaced h_i with the maximum allowed neighbor distance, for example). If adaptive gravitational softening is used for collisionless

[5] For reference, with our definitions, $\Delta t < 0.2\,(h_i/|\mathbf{a}_i|)^{1/2}$ is equivalent to setting the parameter "ErrTolIntAccuracy" ≈ 0.01 in GADGET-2.

[6] Some physics, such as cooling, photo-ionization, and recombination, are handled in a fully-implicit numerical scheme, which (in the limit where, say, the cooling time is much shorter than the timestep) iteratively solves for the equilibrium temperature balancing all heating and cooling physics over each timestep. This means they do not impose an additional explicit timestep criterion.

(star and dark matter) particles, they must obey additional Courant-like timestep criteria given in § 4.3.3.

For star particles, we additionally impose a restriction $\Delta t < \mathrm{MAX}(10^4\,\mathrm{yr},\, t_*/300)$, where t_* is the age of the star; this prevents the code from "skipping" any significant portion of stellar evolution if, somehow, a star formed in a region where the other timestep criteria allowed long timesteps (although this is very rare), and also ensures that the expectation value of the number of SNe per particle per timestep is always < 1 at our production resolution.

Readers interested in more details of the time integration scheme should consult the public GIZMO source code.

2.8 Initial Conditions

All simulations in this paper are fully cosmological "zoom-in" simulations: a large box is simulated at low resolution to $z = 0$, and then the mass within and around the halo(s) of interest is identified, traced back to the starting redshift, and the Lagrangian region containing this mass is re-initialized at much higher resolution for the ultimate simulation (Porter 1985; Katz & White 1993). The initial conditions are generated with the MUSIC code (Hahn & Abel 2011), using second-order Lagrangian perturbation theory to evolve the initial conditions to redshift $z \sim 100$, at which point the GIZMO simulation begins. In the simulations here, the Lagrangian high-resolution regions are defined by a convex hull including all particles within $\sim 5\,R_{\rm vir}$ of the final ($z = 0$) "primary" galaxy (most massive galaxy within the high-resolution region); we have used a series of re-simulations with progressively higher resolution, including baryons, to refine the Lagrangian regions more accurately, with a target of zero low-resolution DM particles contaminating the region within $\sim 2\,R_{\rm vir}$. Typically, these regions include a number of smaller galaxies; however, in this paper, we exclude *any* galaxy with $> 1\%$ contamination by mass within $R_{\rm vir}$.

Table 1 describes the initial conditions for the initial set of halos we have simulated to $z = 0$. We consider a series of halos with different masses; many of these are chosen to match the halos from our FIRE-1 studies (specifically simulations first presented in Hopkins et al. 2014 and Chan et al. 2015). In all cases, the ICs are re-generated if needed to meet our strict contamination standard above. A couple of FIRE-1 ICs are not re-simulated here, because they were not generated from the MUSIC code (they were taken from older work); for consistency and clarity we will only include ICs generated in a consistent manner here. We have added new simulations here to increase our statistical sampling of halo growth histories and mass. The specific halos we re-simulate are chosen to represent a broad mass range and be "typical" in most properties (e.g. sizes, formation times, and merger histories) relative to other halos of the same $z = 0$ mass. Simulations labeled "q" (e.g. **m10q**) have more "quiescent" halo growth histories at late times (i.e. tend to form earlier) while those labeled "v" have more "violent" late-time histories (tend to form later); however we stress that these all lie well within the typical scatter in such histories at each mass (for example, each "q" history has several major mergers at high redshifts). Other labels ("i", "f") are purely for bookkeeping. Several ICs (**m10q**, **m10v**, **m11q**, **m11v**, **m12q**, **m12i**) are taken from the AGORA project (Kim et al. 2013b), to enable easy comparisons with a wide range of different codes (for example, our gravity solver and force-softening prescription has already been tested against other codes for some of these ICs therein). We adopt a standard, flat ΛCDM cosmology with $h \approx 0.70$, $\Omega_{\rm M} = 1 - \Omega_\Lambda \approx 0.27$, and $\Omega_b \approx 0.045$ (consistent with current constraints; see Planck Collaboration et al. 2013).[7]

We scale the resolution with simulation mass, to achieve the optimal possible mass and force resolution for each halo; we study both mass and force resolution extensively in § 4.

For readers interested in more details, or reproducing our results, all initial conditions used here are publicly available.[8]

2.9 Parallelization & Runtime Requirements

The simulations here use a hybrid OpenMP-MPI parallelization scheme with a number of optimizations specific for "zoom-in" simulations. These are listed in Appendix G, together with explicit strong and weak scaling tests. Our improvements allow us to extend good weak scaling on production zoom-in simulations to at least $> 16,000$ CPU cores (and $> 10^6$ cores for large-volume simulations).

With these optimizations, each high-resolution, production-quality simulations of a MW-mass galaxy with particle masses $\sim 7000\,M_\odot$ (a few $\times 10^8$ total particles) typically requires $\sim 10^6$ cpu-hrs; for our smallest dwarfs (particle masses $\sim 250\,M_\odot$), their much lower star formation efficiencies and baryonic densities reduce this to $\sim 10^4$ cpu-hrs. All simulations here were run on the XSEDE Stampede, Comet or NASA Pleiades machines.

Details of our code optimizations and scaling tests are presented in Appendix G.

2.10 A Complete List of Differences Between FIRE-1 & FIRE-2

Although they are discussed in great detail throughout this paper, for the sake of clarity we here summarize the differences between the FIRE-1 and FIRE-2 simulations, in order from most to least important.

(i) **More Accurate Hydrodynamic Solver:** As described in § 2.1, FIRE-2 uses the newer, more accurate mesh-free finite-volume Godunov-type MFM method to solve the hydrodynamic equations. FIRE-1 used the older "pressure-energy" SPH ("P-SPH") method. In § 5 we show how this affects our results; while the differences are generally second-order, this appears to be the single change with the largest effects on our predictions.

(ii) **Manifestly-Conservative Supernovae Ejecta Distribution:** In Appendix D and Paper II, we describe in explicit detail how, algorithmically, we distribute the products (mass, metals, momentum, and energy) of mechanical feedback (SNe and continuous stellar mass loss) from star particles into the surrounding gas particles. As discussed there, the FIRE-1 runs used a slightly simpler algorithm, which can, in situations where the gas elements surrounding a star are highly disordered, produce a distorted (anisotropic) deposition (e.g. biasing the momentum deposition so it is not deposited symmetrically in the rest frame of the star, violating linear momentum conservation). We stress that the FIRE-1 algorithm still ensured the mass and energy of ejecta were exactly conserved; the issue comes with the spatial/vector distribution of the ejecta (momentum conservation). We have developed a novel scheme in FIRE-2 which eliminates this numerical bias and ensures manifest

[7] For the sake of comparison with other work, some ICs are matched to simulations which adopted very slightly different cosmological parameters. These differences are at the $\sim 1\%$ level and their effects are much smaller than standard halo-to-halo variation.

[8] For the MUSIC files necessary to generate all ICs here, see: http://www.tapir.caltech.edu/~phopkins/publicICs

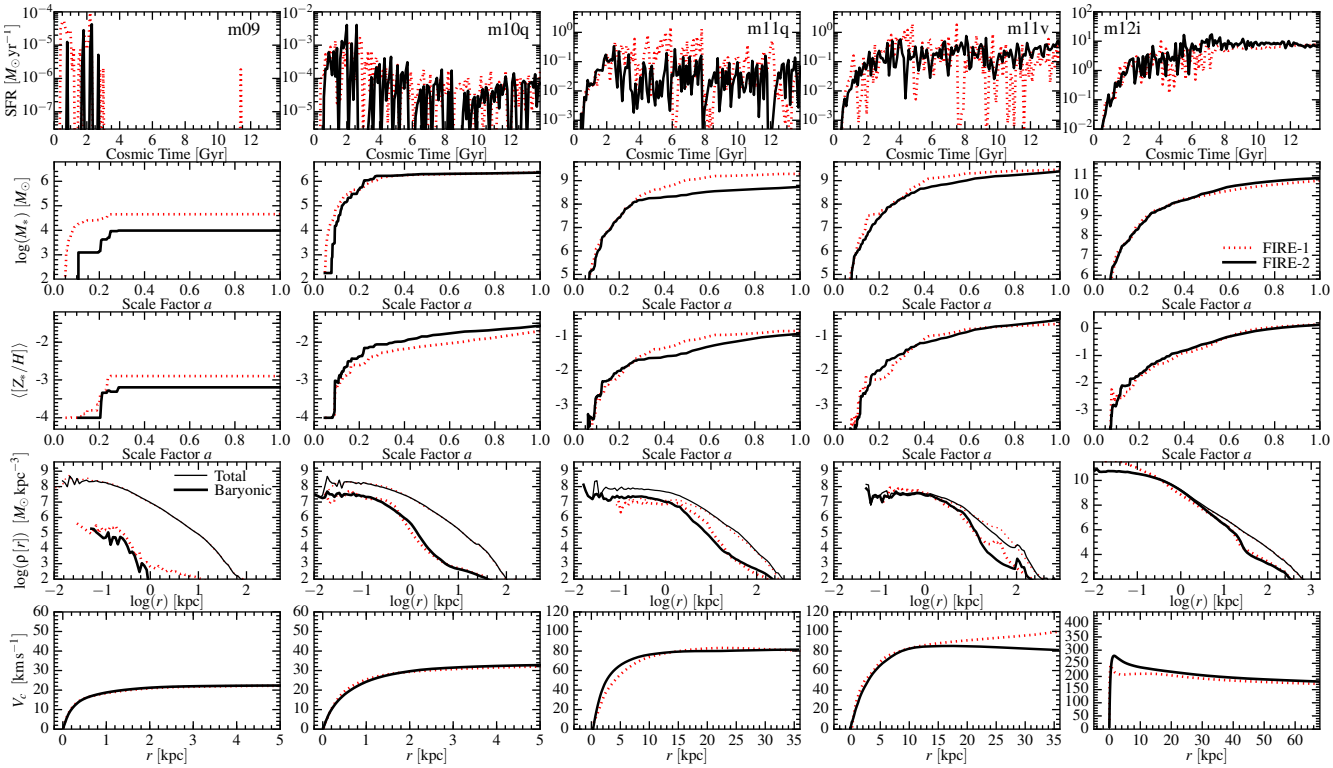

Figure 5. Comparison of galaxy properties and formation histories in FIRE-1 versus FIRE-2, as in Fig. 4. We show galaxies for which the *identical* halo is a member of the "core set" of both FIRE-1 and FIRE-2 simulations. FIRE-2 combines a more accurate hydrodynamic method, higher resolution, a more accurate numerical algorithm for depositing supernova ejecta into gas around explosions, and updated cooling tables (for a complete list of changes, see § 2.10). Nevertheless the results are qualitatively identical in every property that we examine here. We do see some quantitative differences. For dwarf galaxies, we find slightly lower stellar masses, because of the updated photo-heating tables. Massive galaxies show slightly higher masses and central rotation velocities, because of enhanced mixing, which occurs because our more accurate hydrodynamic method changes the cooling and efficiency of wind escape in "hot halos" at late times. The enhanced "burstiness" in FIRE-1 **m11v** occurs because it was run with $\sim 10\times$ lower resolution as compared to FIRE-2. We examine each of the numerical aspects of the method in detail below.

conservation. This difference generally has very small effects, but does appear to influence the central stellar masses/densities of massive galaxies, and because the error term in the older implementation was resolution-independent, it actually can influence galaxy morphologies more significantly at the highest resolutions.

(iii) **More Accurate Photo-Ionization Heating:** In Appendix E and Paper III, we describe our treatment of radiation transport in explicit detail, including UV/optical/IR radiation pressure, photo-electric heating, and photo-ionization heating. The method and source terms are almost entirely identical between FIRE-1 and FIRE-2. In our default treatment of photo-ionization heating specifically, we conduct a gas neighbor search around each star particle, consuming ionizing photons using a local Stromgren approximation until the photon budget is exhausted. However in FIRE-1, the search was simply terminated at the boundary of the local computational domain – any remaining photons were lost. In FIRE-2, any remaining photons are propagated via the long-range tree-based radiative transfer method in Appendix E. The fraction of photons affected is small since the vast majority are absorbed locally, and so this produces weak or negligible differences on galactic scales (nearly undetectable except in small dwarfs), but it eliminates the explicit domain-dependence of local HII regions.

(iv) **Removal of Spurious "Artificial Pressure" Terms:** In FIRE-1, we included an artificial numerical "pressure floor" for cold gas in the ISM, following the approach in e.g. Robertson & Kravtsov (2008). These floors are designed to artificially suppress gravitational collapse of any gas resolved with less than ~ 4 thermal Jeans lengths – artificially ensuring the Truelove et al. (1997) criterion is always met. It has since become clear that this term is (a) redundant, since by definition any regime in which it is significant is one which our star formation/sink particle criteria should convert into stars – essentially, we are "double-counting" sub-grid treatments for un-resolved fragmentation, (b) numerically problematic, when coupled to a self-gravity criterion for star formation, (c) noisy, generating spurious temperature noise at high densities and low temperatures where turbulent, not artificial thermal pressure, actually dominates, and (d) unphysical and often erroneous (it fails to conserve energy and suppresses real fragmentation, especially in turbulence-dominated regions, while providing no visible numerical benefit). We therefore include no such terms in FIRE-2, but instead follow standard practice in the star formation community and rely on the sink-particle (star formation) criterion to treat unresolved fragmentation (see Federrath et al. 2010, for discussion). In § 6 we show the removal of these terms has no effect except to eliminate some obviously unphysical resolution-scale artifacts in the cold gas, as expected.

(v) **Updated Cooling Tables & SNeII Yields:** The physical mechanisms of stellar feedback, and assumptions about stellar evolution, are the same between FIRE-1 and FIRE-2. This means SNe rates (Ia and II), wind mass loss rates and kinetic luminosities, bolometric luminosities and luminosities in different bands, yields, etc., are the same. We have made one minor update: in FIRE-1, we

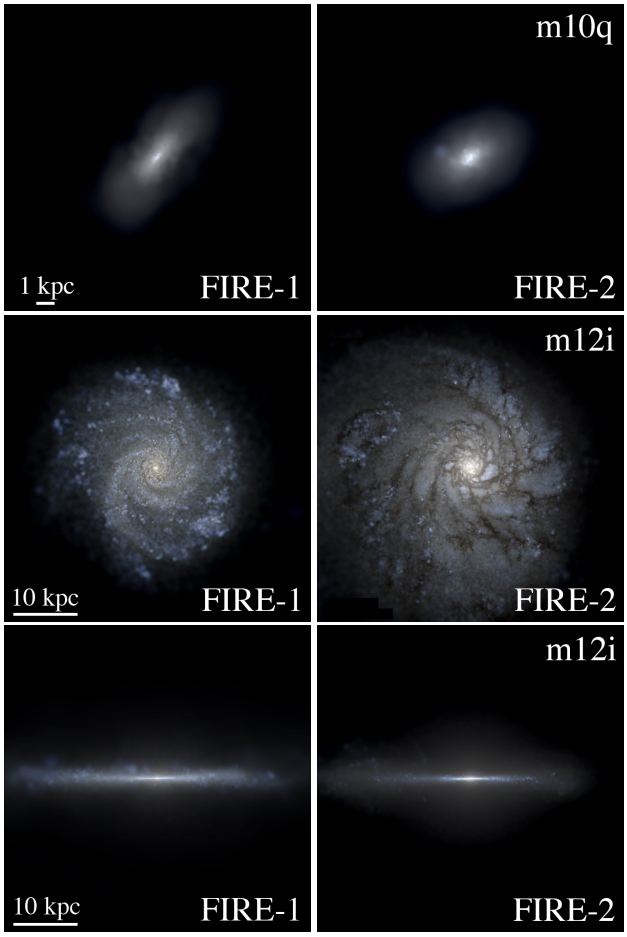

Figure 6. Mock images, as in Fig. 1, comparing FIRE-1 (*left*) and FIRE-2 (*right*) versions of the same galaxy on the same scale. *Top:* Dwarf galaxy (**m10q**). Because their morphologies are irregular or spheroidal, they are insensitive to numerical details. *Middle & Bottom:* MW-mass galaxy (**m12i**) seen face-on and edge-on. Qualitatively, the morphologies are similar. The FIRE-2 run is higher-resolution, which translates to a slightly thinner thin disk and a more extended, low surface-brightness outer disk.

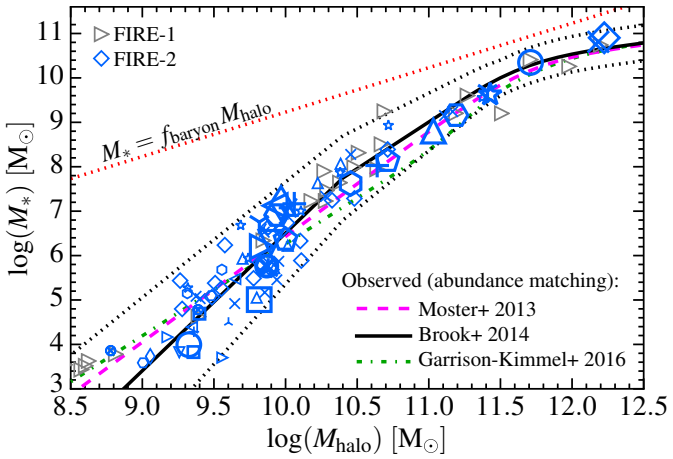

Figure 7. Stellar mass-halo mass relation for FIRE-2 simulations (colored points) at $z = 0$. Stellar masses and halo virial masses are defined as in Table 1, for all resolved, uncontaminated halos (116 galaxies total; see text, § 3). Large points show the "primary" (most massive) galaxy within the zoom-in region, in each simulation (different point styles). Grey triangles show FIRE-1 simulations. While individual galaxies may differ in mass, the effects are primarily stochastic: the two agree well on average. We compare observational estimates as labeled; black dotted lines show the observationally estimated $\sim 95\%$ intrinsic scatter (see text). Within the scatter and systematic variations between fits, the simulations agree well with the observations at all masses.

used the SNe II yields of Woosley & Weaver (1995); however, it is widely known that these older models significantly under-predict the observed yields in Mg and Ne, and we confirmed this in Ma et al. (2016a). We have therefore updated this to the more recent Nomoto et al. (2006) yields, which remedies this issue. We stress, though, that for all other species (especially C and O, which constitute most of the metal mass and are the dominant coolants), the IMF-averaged yield (which is all that appears in our code) is within $\sim 10\%$ of Woosley & Weaver (1995). Since Mg and Ne are negligible coolants, this has no detectable effect on our main results. Similarly, the cooling physics is the same in FIRE-1 and FIRE-2. However we have updated some of the actual fitting functions used to compute the cooling functions (specifically for the recombination rates, photo-electric heating including PAHs, optically-thick cooling, and dust cooling), to match more accurate cooling tables made public since FIRE-1 was developed. For the sake of transparency and clarity, a *complete* set of fits to the FIRE-2 stellar evolution, yield, and cooling tabulations are presented in Appendices A-B.

(vi) **Code Optimization, Higher Resolution:** For FIRE-2, we have made a number of purely numerical optimizations to the GIZMO code, to improve speed and parallelization efficiency (for details, see Appendix G). We have also re-compiled some lookup tables and re-fit cooling functions for greater precision. This has no effect whatsoever on our results, of course, but it has allowed us to run new simulations at even higher resolution compared to FIRE-1, as part of the "Latte" runs (Wetzel et al. 2016). § 4 extensively discusses the effects of improved spatial and mass resolution in the simulations.

3 BASIC RESULTS & COMPARISON BETWEEN FIRE-1 & FIRE-2

Table 1 summarizes all the production-quality FIRE-2 simulations that have been run as of writing this paper. For each, we give the halo virial mass, virial radius, stellar mass of the "target" galaxy (the galaxy used to identify the initial zoom-in region), half-mass radius of the target galaxy, mass resolution of the simulation, and some values describing the "spatial resolution" (because our simulations are Lagrangian, mass resolution is well-defined, but "spatial resolution" is inherently variable: we discuss this in detail in § 4.2). We have considered simulations spanning a $z = 0$ halo mass range from $M_{halo} \sim 10^9 - 10^{12}$, similar to our FIRE-1 simulations. All the simulations here have been run to redshift $z = 0$.

Fig. 1 shows both face-on and edge-on images of two of our FIRE-2 MW-mass systems (**m12i** and **m12f**), at the highest resolution we have studied ($m_{i,1000} = 7$). These use STARBURST99 (in fact, the same assumptions used *in-code*) to compute the stellar spectra as a function of age and metallicity for each star particle, and then ray-trace through the ISM assuming a constant dust-to-metals ratio and physical dust opacities to volume-render the observed images in each band, which we use to construct a mock HST $u/g/r$ composite image as seen by a distant observer. Fig. 2 shows images from within the galaxy: we select a random star ~ 10 kpc from the galactic center and construct a Galactic Aitoff projection of the ray-traced image from all stars in the galaxy to the mock observer. Fig. 3 shows images of several dwarf galaxies from the ultra-faint through LMC mass scales.

Fig. 4 shows several properties of a representative subset of our simulations: the star formation rate and stellar mass versus time (archeological formation history of stars within the $z = 0$ galaxy); the stellar mass-weighted mean metallicity of those stars versus time; the $z = 0$ baryonic and total mass profiles; and the $z = 0$ circular velocity curve. Each property is measured for the "target" galaxy in the simulation. Essentially all of our high-resolution simulations show qualitative behavior similar to one of the galaxies plotted.

Fig. 5 compares the galaxies for which we have both production-quality FIRE-1 and FIRE-2 simulations, specifically. Fig. 6 compares the visual morphology of the same galaxies. Here we can directly compare formation histories and morphologies of the same galaxy, with our improved numerical methods.

In Fig. 7, we plot the stellar mass-halo mass relation for our FIRE-1 and FIRE-2 simulations, compared to observations. We identify all resolved, un-contaminated halos in the high-resolution region and plot their virial masses and the stellar mass of the *central* galaxy in each halo,[9] which is (as expected) always smaller than the total mass in $R_{\rm vir}$ plotted in Fig. 4. The FIRE-1 results here are taken directly from Hopkins et al. (2014). We compare these predictions to recent observational constraints, from a combination of abundance matching and mass modeling. The observational fit from Moster et al. (2013), only includes galaxies with $M_* \gtrsim 10^9\,M_\odot$, but manages to reproduce well the observed local group dwarf luminosity functions to $\sim 10^4\,M_\odot$ (Garrison-Kimmel et al. 2016), so we consider it as well. Brook et al. (2014) combine the constraints from Behroozi et al. (2012) at high masses ($M_* > 10^9\,M_\odot$) with local group and field dwarf constraints. Garrison-Kimmel et al. (2016) perform a similar exercise, allowing the scatter below $M_{\rm halo} < 10^{11}\,M_\odot$ to vary; we plot their best-fit median relations for a constant scatter below this mass ≈ 1 dex (but note that for any scatter in the range $0.5 - 2$ dex, the results are similar). We of course expect the scatter to vary continuously with mass, so we show the 95% inclusion contour if we take the model from Garrison-Kimmel et al. (2016) where the scatter is constant at 0.2 dex above $M_{\rm halo} > 10^{11.5}\,M_\odot$ (the value favored by Behroozi et al. 2012 and Moster et al. 2013), and varies linearly with halo mass as $\sigma_{\rm dex} = 0.2 - 0.2\log_{10}(M_{\rm halo}/10^{11.5}\,M_\odot)$ at lower masses (rising to ≈ 0.5 dex at $M_{\rm halo} \approx 10^{10}\,M_\odot$).

Based on Figs. 1-7, we confirm the conclusions of our previous FIRE-1 studies. In fact (although it is impossible to be exhaustive at this point), we have yet to identify any area in which the FIRE-2 predictions differ qualitatively from FIRE-1 predictions. In future work, we will examine detailed properties of the CGM (e.g. column density distributions of different absorbers) where the hydrodynamic solver could, in principle, have a larger effect.

As in FIRE-1, in FIRE-2 the dwarfs tend to have spheroidal morphologies, with relatively little rotation (see Wheeler et al. 2015b). Especially around $M_{\rm halo} \sim 10^{11}\,M_\odot$, repeated bursts of star formation driving cycles of outflow, subsequent infall, and repeated star formation leads to "puffing up" of the dark matter and stellar orbits, generating large cores in the dark matter profiles (Oñorbe et al. 2015; Chan et al. 2015). This also leads to expansion of the galaxy size and low surface brightness in their stellar distribution (El-Badry et al. 2016). "Bursty" star formation predominates in low-mass dwarfs and high-redshift progenitors of massive galaxies (Sparre et al. 2015; Faucher-Giguère 2017), and in galactic nuclei (Torrey et al. 2016). In FIRE-1 and FIRE-2, MW-mass galaxies develop a clear disk+bulge morphology, with large *thin* stellar disks (despite the presence of strong feedback; Ma et al. 2016d), with clear spiral structure, pronounced radial metallicity and age gradients (Ma et al. 2016b). The stellar (and gas phase) metallicities agree well with the observed mass-metallicity relation both as a function of stellar mass at $z = 0$ and as a function of redshift (Ma et al. 2016a); to the extent that galaxies differ in metallicity between FIRE-1 and FIRE-2, they primarily move *along* the stellar mass-metallicity relationship. In both sets of simulations, galaxies drive strong winds, with higher mass-loading in low mass galaxies, sufficient to place galaxies on the observed relationship between stellar mass and halo mass (Hopkins et al. 2014; Muratov et al. 2015; Hayward & Hopkins 2017). Initial examination of our FIRE-2 runs shows these winds produce covering factors of neutral hydrogen similar to our FIRE-1 simulations, similar to observations at a wide range of galaxy masses (Faucher-Giguère et al. 2015, 2016), although this will be studied in more detail in the future.

There are some modest quantitative differences between FIRE-2 and FIRE-1; most of this manuscript will explore the origin of these differences, but they appear to primarily owe to the change in the hydrodynamic solver. On average, dwarfs are slightly lower-mass in FIRE-2; this owes to both hydrodynamics and the fact that (unlike in FIRE-1) we do not artificially "ignore" ionizing photons when they pass outside numerical domains (hence they can still heat gas). The difference in Fig. 5 can be as large as a factor of ~ 3 for a single galaxy, but this is largely stochastic (since the star formation histories are dominated by a few bursts, a small perturbation to the formation history or feedback strength can lead to factor ~ 2 changes in mass) – Fig. 7 makes this clear, as the *systematic* offset between FIRE-1 and FIRE-2 appears to be a factor < 2, well within the systematic uncertainties in the $M_* - M_{\rm halo}$ relation at $M_{\rm halo} \leq 10^{11}\,M_\odot$. We have also examined this relation at $z = (0.5, 1, 2, 4, 6)$ and find similar agreement, so we refer to Hopkins et al. (2014) for an extensive analysis. Metallicities at the same mass are slightly higher, by $< 30\%$; this owes primarily to updated yield tables; this is far smaller than the factor of $\sim 2 - 5$ systematic uncertainty in the calibration of observed galaxy metallicities. For MW-mass galaxies, the stellar masses are slightly higher in FIRE-2, and the bulges slightly more concentrated (the rotation curves

[9] We use the HOP halo finder (Eisenstein & Hut 1998) to identify halos in Fig. 7, for the sake of consistency with the FIRE-1 results published in Hopkins et al. (2014). This combines an iterative overdensity identification with a saddle density threshold criterion to merge subhalos and overlapping halos. We define halo mass $M_{\rm vir}$ and radius $R_{\rm vir}$ as the Bryan & Norman (1998) virial mass/radius. We discard any halo outside the fully high-resolution region ($> 1\%$ contamination by mass, from low-resolution particles), as well as unresolved halos (with $< 5\times10^4$ DM, < 100 baryonic, or < 10 star particles), and subhalos (any halo within $< 2\,R_{\rm vir}$ of a more massive halo center). The exact value of these cuts makes no difference to our conclusions (comparing $< 0.01 - 10\%$ contamination, minimum $< 100 - 10^6$ DM particles or $1 - 10^4$ baryonic particles, or removing our subhalo cut, gives the same conclusions). We define central stellar mass as in Table 1 iteratively by first measuring the half-mass radius of all stars within a large cut (inside 15% of $R_{\rm vir}$), then taking all stars within $3\times$ this radius (and then re-defining the half-stellar mass radius on these stars). This eliminates all obvious satellites we identify by visual inspection and gives results reasonably close to fitting mass profiles of the central system (a detailed mock observational study will be presented in future work, but for an exponential disk this recovers 97% of the mass). Using a simpler cut of all stars at $< 0.1\,R_{\rm vir}$ gives very similar results, except one case with a $z \approx 0$ merger where our iterative cut more accurately separates the galaxies; but for massive galaxies the $0.1\,R_{\rm vir}$ cut includes stars that are clearly part of the halo (for **m12f**, $0.1\,R_{\rm vir} \sim 30$ kpc, and spreading even as much as $\sim 10^{11}\,M_\odot$ in stars within this radius would give a surface brightness fainter than $\mu \sim 26\,{\rm mag\,arcsec^{-2}}$).

have stronger peaks, by a modest amount). We will show that this is a direct consequence of the hydrodynamic treatment, but is also sensitive to simulation resolution and the SNe coupling algorithm.

We have performed a limited, preliminary analysis of quantities such as mass outflow rates and covering factors of neutral hydrogen in the CGM, which suggest they are also similar in FIRE-1 and FIRE-2. As noted above, however, these will be the subject of a more detailed future study.

4 RESOLUTION IN THE FIRE-2 SIMULATIONS

We now discuss our mass, spatial, and time resolution. For each, we will present a series of tests, and summarize these with a set of resolution criteria.

4.1 Mass Resolution

In Lagrangian or N-body methods such as ours, there is a well defined mass-resolution given by the particle mass m_i in the high-resolution Lagrangian region.

To maintain a well-defined mass resolution scale and minimize N-body integration errors, single gas particles are converted into single star particles with the same masses (rather than, for example, spawning star particles with much smaller/larger masses). Mass loss from stars to gas in O-star/AGB winds and SNe means that particles will not have perfectly equal masses, however, so to prevent pathological behavior in very rare circumstances (if e.g. a single gas particle sees many SNe that increase its mass) we use the standard particle splitting and merging routine from Hopkins (2015) to ensure no particle ever deviates by more than a factor of 3 from the median particle mass. This affects only a tiny number of particles (one in $\sim 10^6$). Averaged over our runs at $z=0$, $\sim 90\%$ (99%, 99.99%) of all particles are within < 0.005 (0.1, 0.2) dex of the median particle mass. We will therefore refer only to the median baryonic particle mass m_i in this paper.

Dark matter particles are always more massive by the universal baryon fraction, $m_{\rm dm} \approx 5\,m_i$; this ensures halos and baryonic galaxies are comparably resolved at initial collapse. Of course, since many galaxies retain only a small fraction of their baryons, and dark matter does not cluster on small scales, dark matter structures tend to be vastly larger than baryonic structures and so are far better mass-resolved.

We define, for convenience, the baryonic particle mass in units of $1000\,M_\odot$, and note that this specifies both baryonic and dark matter particle masses:

$$m_{i,\,1000} \equiv \frac{m_i^{\rm baryonic}}{1000\,M_\odot} \quad (1)$$

$$m_{\rm DM} = \frac{\Omega_m - \Omega_b}{\Omega_b}\,m_i^{\rm baryonic} \approx 5000\,M_\odot\,m_{i,\,1000} \quad (2)$$

The low-resolution regions of the box, outside several virial radii of the main galaxies, are populated by lower-resolution collisionless particles, stepping up in a powers-of-eight hierarchy. This is sufficient for resolution of long-range tidal forces from these regions but we do not consider any halo contaminated ($> 1\%$ by mass) by these particles to be "resolved."

4.1.1 Requirements for "Resolved" Self-Gravitating Structures

The mass resolution "required" to accurately model different phenomena depends, of course, on the question being asked and desired level of accuracy. A wide range of studies have shown that structures with masses of $\sim 5-100$ times the element mass are "believable" in the sense that they exist and can be identified as self-gravitating in higher-resolution re-simulations (Klypin et al. 1999; Springel et al. 2001; Kravtsov et al. 2004; Nurmi et al. 2006; Boylan-Kolchin et al. 2009; Wetzel & White 2010). This corresponds to halos, stellar galaxies, or gas clouds of mass:

$$M_{\rm halo}^{\rm min} \sim 0.6\times10^5\,M_\odot\,N_{10}\,m_{i,\,1000} \quad (3)$$

$$M_*^{\rm min} = M_{\rm cloud}^{\rm min} \sim 1\times10^4\,M_\odot\,N_{10}\,m_{i,\,1000} \quad (4)$$

where $m_{i,\,1000} \equiv m_i/1000\,M_\odot$ is the *baryonic* particle mass, to which we reference all quantities, and $N_{10} = N_{\rm desired}/10$ reflects the "desired" number of particles ($N_{\rm desired}$).

We will show below for both DM halos and subhalos, stellar galaxies, and gas clumps/GMCs within galaxies, only a few particles are sufficient for "converged" results, in terms of masses and mass functions – convergence on other predictions is more demanding.

Of course, the presence of even a few baryonic particles might require a quite massive halo. The observed tight relationship between galaxy stellar mass and halo mass implies that galaxies with $M_* \sim 10^4\,M_\odot$ typically live in halos of mass $M_{\rm halo} \sim 10^9\,M_\odot$; using the $M_* - M_{\rm halo}$ relation at low masses from Moster et al. (2013), we estimate the minimum halo mass with a "believable" baryonic relic is $\sim 10^9\,M_\odot$ at $m_{i,\,1000} \sim 1$. This is actually quite well-resolved in dark matter, with $\gtrsim 10^5$ particles. This also means that $\gtrsim 10^5$ baryonic particles have *participated* in the formation history of the halo and have "cycled through it" (assuming something like the Universal baryon fraction is associated with the halo). It is just that the star formation efficiency is so low that the *residual* mass in stars is small. There is no question, then, that such objects are "real" in the simulations and, if they have such small stellar masses, that feedback had a real effect (it had to prevent $> 10^5$ particles from cooling and forming stars!).

Moreover, because dark matter dominates the gravitational forces in these small galaxies, the orbital dynamics of the surviving stars (determined by the dark matter potential, not the negligible self-gravity) can be believed so long as the dark-matter structure is resolved (i.e. the stars are just tracer particles). This is easily satisfied in halos with $M_{\rm halo} \gtrsim 10^9\,M_\odot$, *independent of the number of stars* (given a realistic $M_* - M_{\rm halo}$ relation).

We will show below that even the *internal* dynamics, evolution over a Hubble time, and mass profile of DM structures, are reliably converged down to radii enclosing just ~ 200 elements. This is much more demanding than "existence" of halos or gaseous structures (which requires just a few elements), as expected.

4.1.2 Requirements for Well-Behaved Internal Hydrodynamics

In Hopkins (2015), we show that, for the MFM method here, a few hundred resolution elements are sufficient to capture the orbital dynamics of thin Keplerian disks for $\sim 10-100$ orbital times (for a galactic disk, roughly equivalent to a Hubble time), the shock structure of strong blastwaves and/or implosions (the Sedov and Noh tests), and all the correct qualitative behaviors of self-gravitating polytropic sphere collapse (the Evrard test) and cosmological structure formation (the Zeldovich test), all to within a factor much better than ~ 2 of the exact solution. This means that a single star particle generating SNe in a halo and blowing out $\sim 100-1000$ surviving gas particles (total baryonic mass $\gtrsim 10^6\,M_\odot\,m_{i,\,1000}$ or, at the Universal baryon fraction, halo mass $\gtrsim 10^7\,M_\odot\,m_{i,\,1000}$) is at least qualitatively well-behaved, if not converged. This is consistent, roughly, with the convergence in properties such as the enrichment history presented below, which require similar particle numbers in the relic baryonic galaxy to resolve the entrainment/recycling of

Table 3. Illustrative Examples of "Resolution" Scales in a MW-Mass Halo (**m12i**)

Property	Notation	Resolution Level Run to $z = 0$			
		$m_{i,\,1000} = 450$	$m_{i,\,1000} = 56$	$m_{i,\,1000} = 7$	$m_{i,\,1000} = 0.88$ (DM-only)
Particle Number	$N_{\rm tot}$	2.5×10^6	2.0×10^7	1.4×10^8	5.6×10^8
Baryonic Particle Mass ($M_\odot$)	m_i	4.4×10^5	5.6×10^4	7070	–
Minimum Timestep (yr)	$\Delta t_{\rm min}$	600	260	120	1000 (no gas)
Star-Forming Densities:					
Minimum Density of Star Formation (cm^{-3})	$n_{\rm SF,\,min}$	100	1000	1000	–
Mean Density of Star Formation (cm^{-3})	$\langle n_{\rm SF}\rangle$	700	1900	3400	–
Gas Resolution (Inter-Particle Separation = Force Softening) at Star-Forming Densities:					
Minimum Spatial Resolution (pc)	$h_i^{\rm min}$	5.0	1.4	0.38	–
Spatial resolution at $n_{\rm SF,\,min}$ (pc)	$h_i^{\rm threshold}$	57	13	7	–
Spatial resolution at $\langle n_{\rm SF}\rangle$ (pc)	$h_i^{\langle SF\rangle}$	30	10	4.6	–
Jeans Scales in Warm (10^4 K) ISM, Corresponding to Marginally-Resolved (10-element) Structures:					
Minimum Jeans Radius $= L_{\rm Jeans}/2$ (pc)	$\lambda^{J,\,WIM}$	130	17	2.1	–
Maximum Density with Resolved $M_{\rm Jeans}$ (cm^{-3})	$n_{\rm max}^{J,\,WIM}$	20	1200	7.4×10^4	–
Turbulent Jeans Scales in Cold Clouds ($T = 10$ K, $\Sigma_{\rm GMC} = 300\,M_\odot\,{\rm pc}^{-2}$), Corresponding to Marginally-Resolved (10-element) Structures:					
Minimum Turbulent Jeans Radius (pc)	$\lambda^{\rm turb,\,CNM}$	42	15	5.3	–
Maximum Density with Resolved $M_{\rm Jeans}^{\rm turb}$ (cm^{-3})	$n_{\rm max}^{\rm turb,\,CNM}$	570	1600	4600	–
Dark Matter Resolution:					
Particle Mass ($M_\odot$)	$m_{\rm DM}$	2.7×10^6	3.4×10^5	4.3×10^4	5400
Minimum Inter-Particle Separation (pc)	$h_i^{\rm DM,\,min}$	66	32	16	8.2
RMS Separation within galaxies at $z = 0$ (pc)	$h_i^{\rm DM,\,core}$	330	150	70	38
DM Convergence Radius, $t_{\rm relax} = 0.06\,t_{\rm circ}(R_{200})$ (pc)	$r_{0.06}$	670	300	150	75
Typical "N-Body Heating" Rates:					
Gas-Gas Scattering (erg cm^3 s^{-1})	$\langle \mathcal{Q}_{\rm heat}^{\rm gas-gas}\rangle$	8×10^{-30}	1×10^{-30}	1×10^{-31}	–
Gas-DM Scattering (erg cm^3 s^{-1})	$\langle \mathcal{Q}_{\rm heat}^{\rm gas-DM}\rangle$	2×10^{-27}	2×10^{-28}	5×10^{-30}	–

Several "spatial resolution" and "time resolution" properties of the simulations discussed in § 4. We focus on our resolution study of the **m12i** system, but results for other MW-mass galaxies are nearly identical, and results for dwarf galaxies are qualitatively similar at similar mass resolution. Because our simulations are Lagrangian, only mass resolution is truly "fixed." Spatial (both hydrodynamic and gravitational, which are the same always) and time resolution are both adaptive, and in principle can reach arbitrarily small values, but in practice reach minimum values based on the densest mass-resolved structures in the simulation.

SNe ejecta and its re-incorporation into subsequent generations of star formation.

4.1.3 Requirements for Convergence In ISM Properties, Star Formation, and Stellar Feedback

The most demanding mass resolution requirements come from the physics of the ISM and CGM.

(i) **The Toomre Mass & ISM Structure:** In a series of previous studies, we showed that reliably capturing star formation (both rates and spatial distribution) requires the ability to resolve at least the *existence and self-gravity* of the largest self-gravitating gas structures (i.e. fragmentation) in a galactic disk (see Hopkins et al. 2011, 2012c,b, 2014). These studies showed that once this criterion is met, the SFR predicted in the simulations becomes independent of the numerical star formation model (the density threshold and assumed rate of conversion of dense, self-gravitating gas into stars), because it is regulated by feedback and the rate-limiting step is the formation of these first fragments. This means we do not strictly need to resolve the internal dynamics of such clouds (only their formation). And although there will always be clouds and sub-structure below our resolution limits, both observations and simulations have repeatedly shown that almost all star formation occurs in the largest GMC complexes in galaxies (see Evans 1999; Rosolowsky et al. 2003; Blitz & Rosolowsky 2005; Bolatto et al. 2008; Elmegreen 2007; Tasker & Tan 2009; Feldmann & Gnedin 2011; Murray 2011; Harper-Clark & Murray 2011; Hopkins 2012a). This mass $M_{\rm GMC}^{\rm max}$ is set by the Toomre mass, $M_{\rm Toomre} \sim f_{\rm gas}^2\,M_{\rm gas}$ where $f_{\rm gas}$ is the gas fraction inside the relevant effective

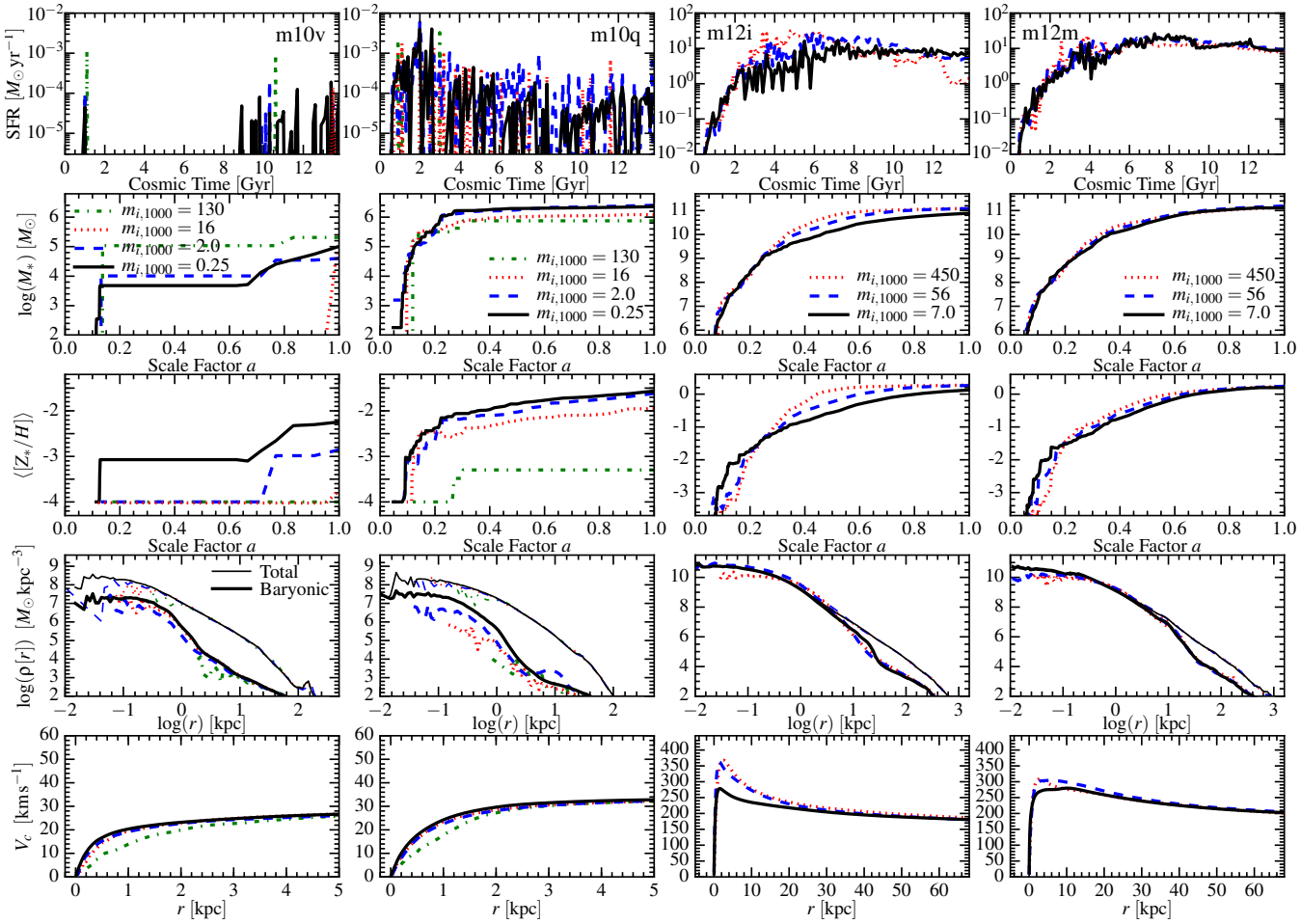

Figure 8. Resolution study, as in Fig. 4. Each column is one galaxy, and each line shows a different mass resolution ($m_{i,1000}$) up to our production resolution. *Top:* Star formation rate versus cosmic time. In dwarfs (**m10v, q**), low resolution leads to artificially large "burstiness": a few bursts dominate the history and eject the remaining baryons. At higher resolution, star formation is still "bursty", but it becomes robust to changes in resolution. In MW-mass galaxies, higher resolution shifts star formation to slighly later times by more efficiently regulating the low-mass projenigtor at high redshifts. *Second from Top:* Stellar mass versus scale factor. Note that the $m_{i,1000} = 16$ (130) simulations of **m10v** have only 10 (2) star particles in the main halo, and yet they are within a factor ~ 3 of the converged stellar mass; by the time ~ 100 particles are in the main galaxy, the mass is converged to a factor ~ 1.2. *Middle:* Average stellar metallicity versus scale factor. This converges more slowly than stellar mass; with < 100 star particles, dwarf galaxies show artificially suppressed metallicity, because low resolution under-samples the enrichment history and leads to artificially bursty SFH that blows out metals completely. Massive galaxies show smaller differences that match differences in their SFHs. *Second from Bottom:* Mass density profile at $z = 0$: convergence is good down to radii enclosing $\sim 100-200$ particles of the "type" of interest (details below). *Bottom:* Circular velocity profile at $z = 0$. For dwarfs, this is dark-matter dominated and therefore under-resolved only at our lowest resolution (where the DM "convergence radius" discussed below is > kpc). For the MW-mass galaxy, the modest difference in the SFH at $z \sim 0.5-2$ in **m12i** translates to a factor ~ 2 difference in the central bulge mass, which leads to a more strongly peaked central V_c at low resolution. **m12f** (not shown) shows nearly identical behavior to **m12i**; **m12m** shows better agreement at all resolution levels.

radius of the baryonic galaxy. So, if we require $M_{\rm cloud}^{\rm min} \lesssim M_{\rm Toomre}$, we obtain the desired resolution criterion $m_{i,1000} \lesssim m_{i,1000}^{\rm Toomre}$ where

$$m_{i,1000}^{\rm Toomre} \lesssim \frac{100}{N_{10}}\,\frac{M_{\rm Toomre}}{10^6\,M_\odot} \sim 100\left(\frac{f_{\rm gas}^2\,M_{\rm gas}}{10^6\,M_\odot\,N_{10}}\right) \quad (5)$$

This defined the target resolution of the original FIRE-1 massive-galaxy simulations ($10^{12}\,M_\odot$ halos).

(ii) **SNe Cooling Radii:** A still more demanding criterion is set by SNe physics. As discussed in detail in Paper II (as well as Cioffi et al. 1988; Thornton et al. 1998; Martizzi et al. 2015; Walch & Naab 2015), a single SNe remnant radiates its thermal energy rapidly upon reaching a nearly-invariant "swept up" mass $M_{\rm cool} \sim 3000\,M_\odot$; if a number $N_{\rm SNe}$ occur before the cooling time expires, the "cooling mass" simply scales $M_{\rm cool} \propto N_{\rm SNe}$. If SNe energy is injected in a kernel-weighted fashion over $N_{\rm NGB}$ elements as we do here, then almost all of the energy goes into the nearest $N_{\rm NGB}^{1/3}$ neighbors; requiring they have a total mass $< M_{\rm cool}$ in turn requires:

$$m_{i,1000}^{\rm SNe} \lesssim 0.9\,(N_{\rm NGB}/32)^{-1/3}\,N_{\rm SNe} \quad (6)$$

This set the target mass resolution for the original FIRE-1 dwarf galaxies ($< 10^{10}\,M_\odot$ halos). As we show in Paper II, unless $m_{i,1000} \ll m_{i,1000}^{\rm SNe}$, it is necessary to properly account for the conversion of energy into momentum in the unresolved Sedov-Taylor (S-T) phases. Failure to do so will significantly under-estimate the effects of feedback.[10]

[10] Note that Eq. 6 does not mean we cannot resolve hot gas and/or overlapping SNe bubbles at lower resolution. If any mass of stars m_0 forms and the corresponding $N_{\rm SNe} \sim m_0/100\,M_\odot$ go off in an overlapping resolution ele-

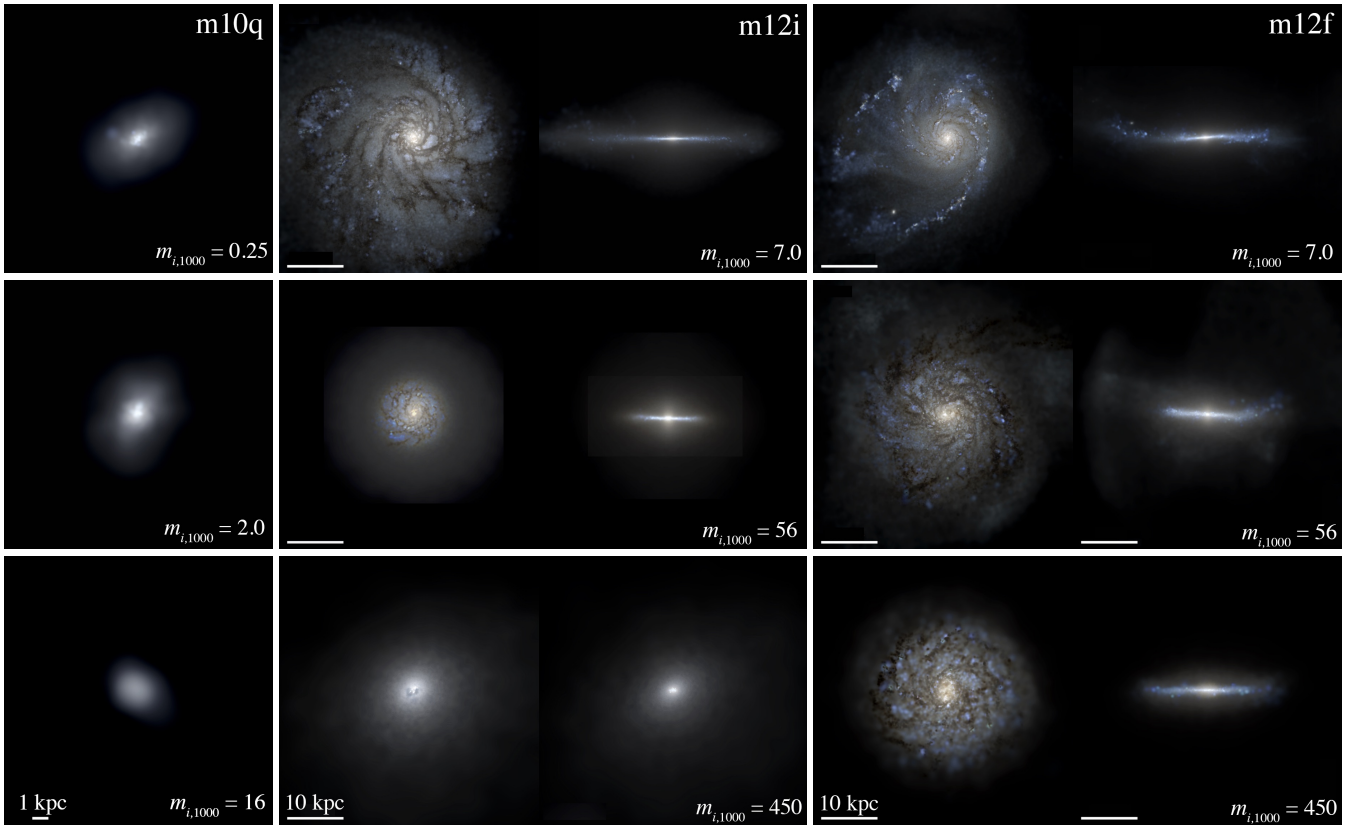

Figure 9. Mock images, as in Fig. 3, showing the effects of resolution on the morphology of the simulated galaxies at $z = 0$, for a subset of the resolution series from Fig. 8. *Left:* Dwarf galaxy (**m10q**): because dwarf morphologies are disordered, their qualitative morphology does not depend sensitively on resolution (this is true for all our dwarfs). *Center:* MW-mass galaxy **m12i**: In **m12i** we see a trend towards a more extended thin disk component dominating as we increase resolution. The effective radius also increases, but much less dramatically (just by $\sim 40\%$ from $m_{i,\,1000} = 56$ to $m_{i,\,1000} = 7.0$); the bulge and central $\sim 2-3$ kpc remain similar in each case, but the extended, gas-rich disk is much more prominent. At the lowest resolution ($m_{i,\,1000} = 450$), we see little disk-like structure at all (although there is a rotating gas+stellar disk ~ 1 kpc in size). At intermediate resolution, a clear disk with ~ 5 kpc radius appears – there is also an extended, smooth-light component out to ~ 10 kpc. At our highest resolution, this extended component exhibits spiral structure and the gas+young stellar disk extends to > 10 kpc from the galaxy center (albeit at low surface brightness: all figures here use an 8-magnitude stretch). *Right:* MW-mass galaxy **m12f**: In this slightly more-massive galaxy, an extended disk is present at all resolution levels (the trends seen in **m12i** are still present, but much weaker). **m12m** (not shown) similarly shows a disk at every resolution level.

(iii) **Dwarf Galaxy "Burstiness":** In previous papers we have shown that for small dwarf galaxies, star formation is robustly "bursty";[11] however, at low resolution, the "burstiness" is artificially enhanced by numerical effects (see e.g. Sparre et al. 2015). This owes to the fact that stars form in units of star particles. At sufficiently low resolution, a single star particle implies a massive, co-eval star cluster, therefore a large number of approximately co-eval SNe. Although we know SNe are clustered in reality, this artificial numerical clustering could easily "overshoot" reality, producing too many synchronized SNe in the same location, which in turn leads to an over-large super-bubble that could heat the entire galaxy gas supply to super-virial temperatures. If we consider the "unit" of star formation (minimum resolved cloud) to be, say ~ 10 particles (since a single lone gas particle cannot be at much higher density than its neighbors), then taking the IMF-average $N_{\rm SNe} \sim m_i/100\,M_\odot$ per particle, the entire baryonic mass in the galaxy can be heated to $T \gg 10^6$ K ("super-bubble" temperatures) if that baryonic mass is below $\sim 600\,m_i$. So resolving "venting" of individual super-bubbles (much smaller than the entire galaxy) requires $\gtrsim 1000$ gas particles in the galaxy. Assuming further a gas fraction of ~ 1 during the gas-rich phases of star formation and/or typical dwarf galaxies, and using the stellar mass-halo mass relation above, this implies a particle mass target:

$$m_{i,\,1000}^{\rm burstiness} \lesssim 5 \left(\frac{M_{\rm halo}}{10^{10}\,M_\odot}\right) \quad (7)$$

(iv) **The Jeans Mass:** A common mis-conception is that one needs to resolve the Jeans mass in order to capture basic fragmentation physics. This is incorrect, for two important reasons. First, the warm/cold ISM is super-sonically turbulent, so the "fragmentation cascade" (hierarchical structure of fragmentation into GMCs, clumps, cores, etc.) is determined by the *turbulent* Jeans mass, not the *thermal* Jeans mass, at least down to the sonic scale

ment within a cooling time, then they can heat a mass $\sim 60\,m_0$ to $> 10^6$ K. So because star formation is clustered, we can still resolve super-bubbles and galactic chimneys at relatively modest resolution. However, resolving the full hot gas content, venting, momentum contribution from confined blastwaves, and early evolution of SNe explosions requires a criterion like Eq. 6.

[11] For the sake of quantitative comparison, we define a specific measure of "burstiness" as the standard deviation in the quantity $\log(\dot{M}_*(\Delta t_1)/\dot{M}_*(\Delta t_0))$ where $\dot{M}_*(\Delta t)$ is the SFR averaged over a time interval Δt, and we compare a short interval $\Delta t_1 = 10$ Myr and longer interval $\Delta t_0 =$ Gyr.

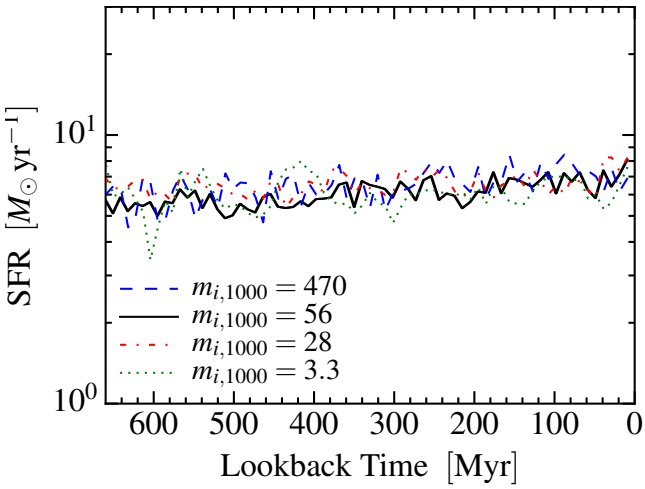

Figure 10. Star formation rate versus lookback time, in our MW-mass **m12i** simulations, re-started at late times. In each case, we re-start using the snapshot at $z = 0.06$ from our run with $m_{i,\,1000} = 56$ (in Fig. 8) as our initial condition, so the late-time ICs of the different realizations are identical. We then run from $z \approx 0.06 - 0$ ($\sim 700\,$Myr physical time), to study how the SFR varies in a massive galaxy given the same initial galaxy properties (that is, factoring out how variations affect the earlier phases when the progenitor was low-mass). Here, we use particle splitting/merging before running to vary the mass resolution. We see that for the same ICs, SFR is well-converged at the resolutions that we consider; variations in the SFR with resolution in Fig. 8 are dominated by (1) early stages when the progenitor galaxy was much lower mass, hence much less well resolved, and (2) less prominently, weak resolution dependence of wind mixing versus escape from the outer halo, which changes the inflow rate back into the galaxy at later times.

($R \lesssim R_{\rm sonic} \sim 0.1\,$pc and mass scale $\sim 1\,M_\odot$). We discuss this below, but the turbulent Jeans mass of a super-sonically turbulent GMC with virial parameter $\alpha \sim 1$ is of order the GMC mass itself ($\sim 10^7\,M_\odot$), whereas the thermal Jeans mass is $\sim 0.1\,M_\odot$. Second, in a homogeneous medium, the Jeans mass defines the *smallest* scales of the fragmentation cascade – by definition, all larger scales are also unstable. Therefore, failure to resolve the Jeans scale simply means all *resolved* scales fragment as they should – the fragmentation cascade is just truncated at the resolution scale, instead of the (smaller) Jeans scale. This is analogous to "resolving turbulence" in the ISM; the actual Kolmogorov ("termination") scale of the turbulent cascade ($\sim$ au) is far smaller than achievable resolution, but that does not mean turbulence cannot be captured (it simply limits the dynamic range of the cascade that can be followed). In both cases (for the same physical reason, in fact; see Hopkins 2012a), the power in the cascade is dominated by the largest-scale structures, so integrated quantities converge very quickly with increasing resolution (Hopkins 2013b; Guszejnov & Hopkins 2015, 2016). We show this explicitly below – the mass function of dense, cold clouds is well behaved regardless of resolution, it simply extends to smaller and smaller sub-structures as we increase the resolution, down to structures of a few particle masses. Thus, the full cascade is not followed, but this is not problematic, provided that the goal of our simulation is not to resolve individual brown dwarf or star formation (the actual "end point" of the fragmentation cascade in the ISM; Hopkins 2012b). This is where the sub-grid model for star formation enters – it is, explicitly, a sub-grid model for the un-resolved fragmentation cascade from a resolved, self-gravitating cold gas clump/cloud all the way down to an aggregate of individual stars with some IMF. The Toomre scale, on the other hand, is important to resolve, because it is the *largest* unstable scale in a disk – in other words, failure to resolve the Toomre scale means that no fragmentation will occur (when it physically should).

4.1.4 *Mass Resolution Tests*

We now justify these resolution criteria, and explore in detail how mass resolution can alter our conclusions. Figs. 8-14 present a series of explicit mass resolution tests of our full-physics FIRE-2 simulations. We consider yet more tests below that focus on the force/spatial resolution. Table 3 gives some typical values for mass, spatial, and time resolution for **m12i** runs at different resolution.

Fig. 8 shows our full-physics cosmological simulations at varying mass resolution. We compare two low-mass dwarfs (**m10v** and **m10q**) and two MW-mass galaxies (**m12i** and **m12f**), to bracket the range of behaviors as a function of mass, and provide some estimate (with two of each) of how robust our conclusions are at a given mass. We compare the same list of properties as Fig. 4. Fig. 9 compares images of the galaxies, to illustrate their visual morphologies.

For the dwarfs, the total stellar mass (and circular velocity profile) converges to better than a factor < 2 remarkably quickly, with just $\sim 2 - 16$ star particles in the $z = 0$ galaxy. This is because the total stellar mass is set by an integral balance between feedback energetics and gravity binding the baryons. Also, given the low particle masses, even these low-resolution dwarf runs easily satisfy our Toomre-mass criterion (Eq. 5). However, at such low resolution ($m_{i,\,1000} = 16 - 130$) they do *not* satisfy our "numerical burstiness" or "SNe cooling" criteria (Eq. 7 and 6, respectively). As a consequence, the SFHs are visibly "too bursty" (dominated by just a couple large bursts), and the metallicities are systematically under-predicted. The latter occurs because the single bursts blow out nearly all the baryons (and metals) from the galaxy: they fail to capture partial entrainment/mixing/incomplete blowout that would keep the metals in the galaxy. The metallicity and burstiness of the SFH appear to converge to factor ~ 2 ($\sim 10\%$) when the number of stars reaches > 100 (> 500) at $m_{i,\,1000} = 2$, as expected from our criterion in Eq. 7. More subtle properties, such as the internal SF structure of bursts and hot gas properties of the galaxy, and escape fraction of ionizing photons (Ma et al. 2015, 2016c), require still higher resolution, hence our highest-resolution runs, with $m_{i,\,1000} = 0.25$, satisfying Eq. 6. These results are also supported by the analysis of dwarf satellite galaxies of our massive, high-resolution **m12i** halo in Wetzel et al. (2016); the convergence in the stellar mass function of dwarfs appears good down to $\sim 5 - 10$ star particles per galaxy; but the metallicities of the satellites with $\lesssim 100$ star particles are suppressed relative to our higher-resolution isolated-dwarf simulations and observations (compare Ma et al. 2016a). Not surprisingly, because the gross morphology of the dwarfs is irregular, it is essentially resolution-independent.

For the massive, MW-mass system, even the lowest-resolution ($m_{i\,1000} = 450$) run has $\gg 10^5$ particles in the baryonic galaxy and easily satisfies the Toomre and "burstiness" criteria (although its progenitor dwarf galaxies at high redshifts may not). As such the SFH and metallicity are more accurately converged. However we do see higher resolution systematically shifts the SFH from early to later times (by a factor ~ 2), as higher resolution allows better resolution of two key physics. First, the generation of winds via resolution of the hot gas "channels" and their escape ("venting") from a multi-phase halo (our "SNe" criterion; for explicit studies see Hopkins et al. 2012b, 2013c; Muratov et al. 2015). At low resolution, hot gas is necessarily "dragged" by a large mass of

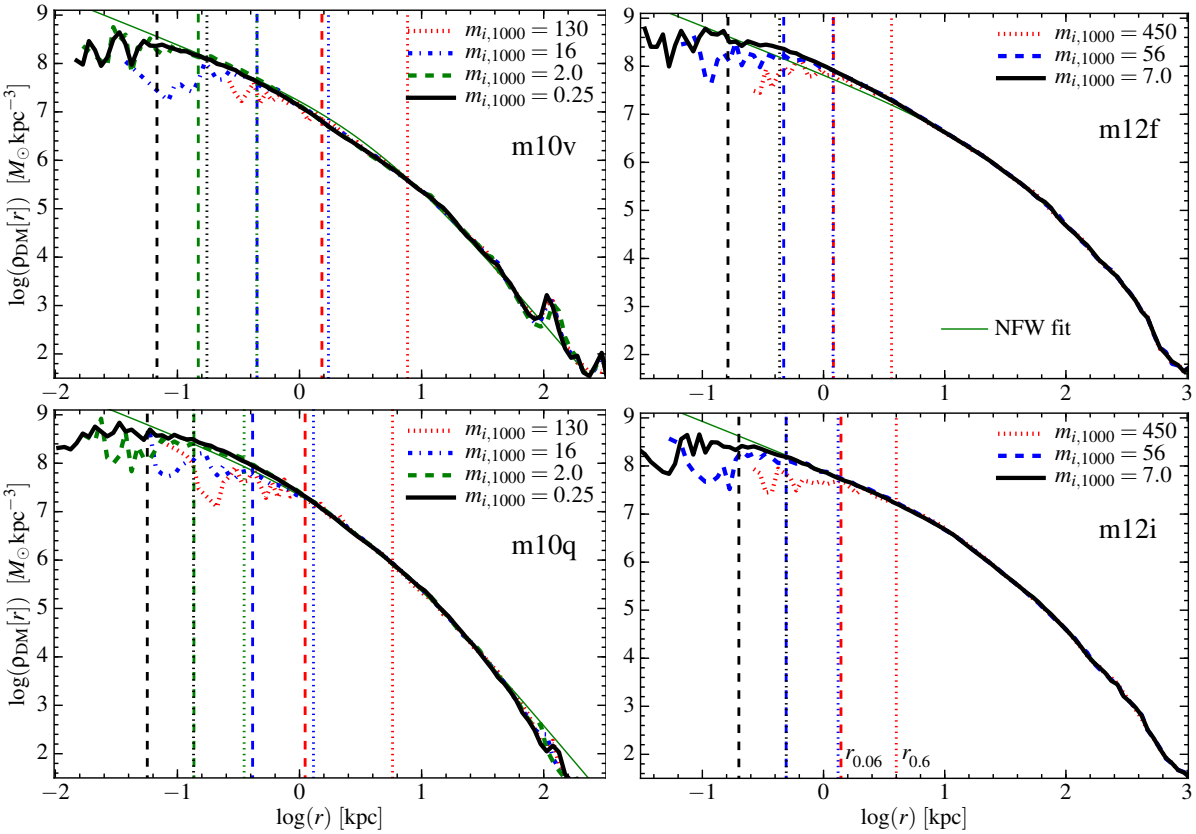

Figure 11. Resolution studies *in DM-only simulations* of the dwarf and MW-mass halos from Fig. 8. We show the radially averaged DM mass profile at $z = 0$ around the main halo, varying the mass resolution. For each factor of 8 in mass resolution, we change the DM force softening by a factor of 2. With increasing mass resolution we converge more and more closely to an NFW-like profile at small radii (thin green line). At radii larger than the Power et al. (2003) radius, $r_{0.6}$. where the relaxation time $t_{\rm relax} \approx 0.6\,t_{\rm circ}(R_{200})$, which encloses approximately ≈ 2200 particles (dotted vertical lines), the convergence is near-perfect. However, even at radius, $r_{0.06}$, where the $t_{\rm relax} \approx 0.06\,t_{\rm circ}(R_{200})$, which encloses just ≈ 200 particles (dashed vertical lines), the convergence is still quite good: densities are under-estimated by at most $\sim 0.05-0.15$ dex. We define the latter as our DM "convergence radius" throughout this paper.

implicitly-coupled gas, because mass is locked into massive particles: at our lowest resolution, for example, a single SNe cannot affect less mass than $\sim 10^6\,M_\odot$ – this means much lower "launch velocities" and temperatures, unless a huge number of SNe explode simultaneously (see Paper II). Second, high resolution allows better resolution of the "burstiness" and SFHs within the smaller *progenitor* galaxies of the more massive $z = 0$ MW-mass system (note that the largest differences appear at early times, when the galaxy is a progenitor dwarf), which reduces their stellar and gas masses, leading to more gas expelled to large enough radii where its recycling times are long. The gas is still re-incorporated eventually, evident in the similar late-time masses and SFRs, but appears to re-accrete later.

As a result of these effects, at higher resolution the final $z = 0$ MW-mass galaxy is slightly lower-mass, but more importantly, because it shifts star formation and re-accretion of recycled material to later times, that material carries larger angular momentum, and the galaxy is less compact. For **m12i**, the effective radius in our high-resolution $m_{i,1000} = 7$ run is ~ 1.4 times larger than the $m_{i,1000} = 450$ run, which translates to a factor ~ 1.8 lower mass inside < 5 kpc; this in turn lowers the peak in the rotation curve from $\sim 370\,{\rm km\,s^{-1}}$ to $\sim 270\,{\rm km\,s^{-1}}$. While the differences are quantitatively modest, the galaxy rotation curve is noticeably "less bulgy" at high resolution, as a consequence. The differences in **m12i** are more dramatically evident in the low surface-brightness outer disk morphology, which goes from being entirely absent (along with any thin disk) at very low resolution to quite prominent at high resolution. Interestingly, however, in **m12f**, although the change in effective radius and the rotation curve with resolution is similar to **m12i**, the visual morphology changes much less dramatically – even at $m_{i,1000} = 450$ there is still a prominent, extended thin disk. This may owe to the fact that **m12f** has a somewhat larger mass (so is proportionally better-resolved at all resolution levels), but likely also owes to the specific fact that **m12i** has a series of mergers around $z \sim 1$, which launch strong fountains and change the angular momentum of the gas that will form its disk (whereas **m12f** grows more smoothly) – so it appears **m12i** is simply more sensitive to resolution effects. Clearly, it is important to push to even higher resolution (in progress), to test whether or not this is fully-converged.

If the dependence on resolution in our MW-mass runs owes to a combination of (1) better-resolving the progenitor (dwarf) galaxies, and (2) resolving channels by which hot gas outflows can escape, we should see similar SFRs independent of resolution at late times *if we start from the same ICs*. In other words, it is useful to test whether indeed the difference with resolution comes from progenitor influence on the late-time galaxy, or whether it is present for fixed conditions in a massive MW-mass galaxy. We therefore consider a series of simulations where we use our **m12i** simulation with $m_{i,1000} = 56$, re-started at $z \approx 0.06$ from a snapshot of the simulation in Fig. 8, but using our particle splitting/merging routine to first split/merge the ICs until they are re-sampled with a desired target

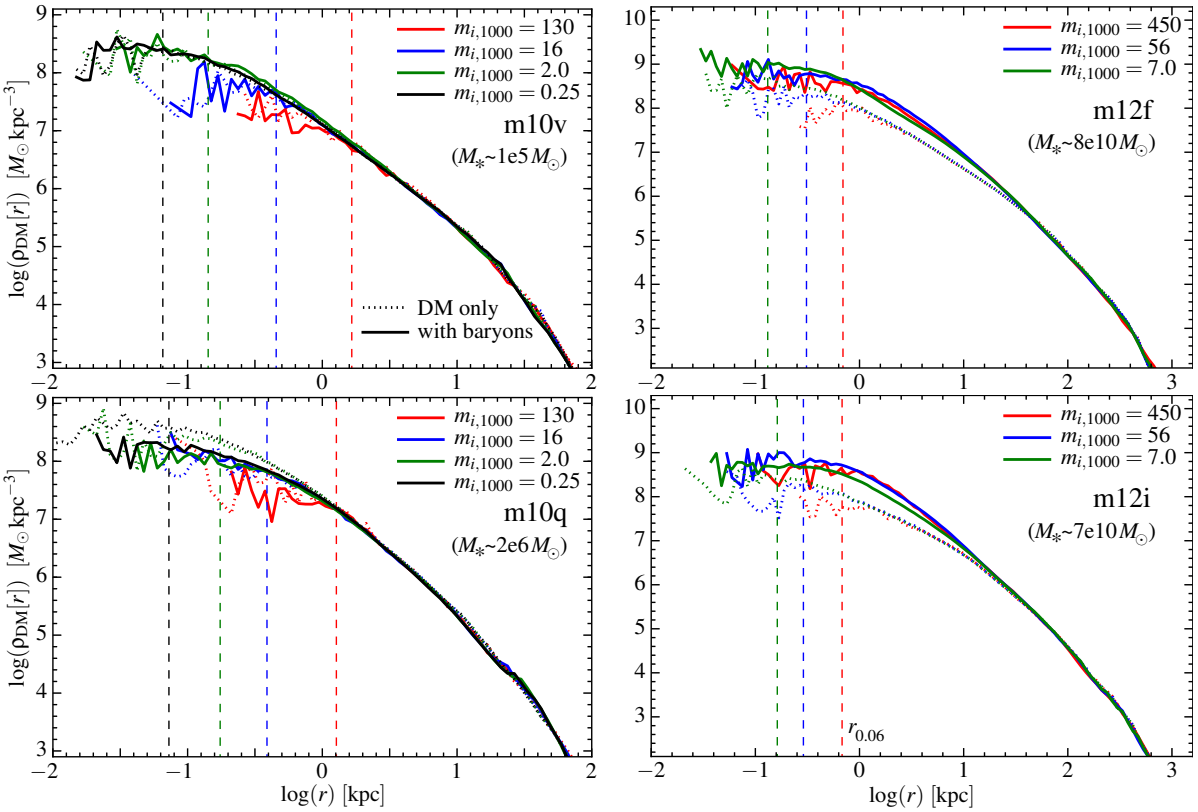

Figure 12. Resolution studies of DM mass profiles in the dwarf and MW-mass halos from Figs. 8-11, in simulations with baryons (solid) compared to DM-only simulations (dashed). We re-normalize the simulations with baryons by $\Omega_m/(\Omega_m - \Omega_b)$ so that if the baryons traced the DM perfectly, the DM-only and baryonic simulations would agree exactly. For clarity, we label the convergence radius ($r_{0.06}$, enclosing ≈ 200 DM particles; vertical dashed lines) for just the simulation with baryons. In **m10v**, the galaxy is sufficiently low mass ($M_* \sim 10^5\,M_\odot$) that it is DM dominated and has little or no core, so DM and baryonic simulations agree well. We also see excellent convergence down to radii ~ 30 pc enclosing just ~ 10 particles. In **m10q**, the more massive dwarf is still DM-dominated but forms a core in the baryonic runs (suppression of $\rho_{\rm DM}$ to ~ 600 pc); the core is actually better-converged than the DM-only runs, down to similar radii ~ 30 pc. In MW-mass runs, the baryons dominate the central mass and cause some contraction of the DM (although less than would be expected from pure adiabatic contraction given the galaxy mass). Differences in the central DM profile are, in these galaxies, dominated by differences in the baryonic mass (high-resolution runs can be *less* dense, owing to slightly smaller stellar masses), not by traditional DM resolution considerations such as N-body relaxation.

resolution. Fig. 10 shows that when we do this, the SFR is nearly identical over ~ 2.5 dex in mass resolution (small deviations at the highest resolution owe mostly to some artifacts of our very aggressive particle splitting/up-sampling routine applied for this specific test).

This is consistent with our argument above, that we only need to marginally resolve the Toomre scale to achieve convergence in the SFR, *given* a specific gas disk initial condition. Quantities such as the Kennicutt-Schmidt relation are thus extremely robust to mass resolution. However, non-linear, long-timescale effects in cosmological simulations (e.g. recycling) shift the SFR *by changing the supply or loss rate of gas in the CGM*.

To understand how much of the resolution dependence owes to purely gravitational physics, Figs. 11-12 consider how the dark matter mass profiles of the galaxies change with mass resolution. First we consider DM-only simulations (i.e. gravity-only simulations), in Fig. 11. We see excellent convergence in the mass profiles of the DM halos (as well as their substructure mass functions, shown below, and halo formation/growth histories). As expected, the $z = 0$ DM mass profiles are well-fit by a Navarro et al. (1996) (NFW)-like profile, down to some minimum scale where numerical effects flatten the profile. Because of the Lagrangian nature of our code, improving mass resolution also improves the effective spatial resolution/force resolution; in fact, we will show below that the nominal spatial force softening is generally much less important than mass resolution. Power et al. (2003) argue that the central "flattening" in DM profiles is dominated by N-body relaxation, and that good convergence should be obtained outside a radius where the N-body relaxation time $t_{\rm relax} \sim 0.6\,t_0$ (where $t_0 \equiv t_{\rm circ}(R_{200}) = 2\pi\,R_{200}/V_{200}$) is comparable to the Hubble time. This is effectively a requirement on the number of particles: $(5/\sqrt{8})\,(N[<r]/\ln N[<r])\,(\rho_{\rm crit}/\bar{\rho}[<r])^{1/2} < 0.6$, satisfied for the densities in Fig. 11 when $N \gtrsim 2200$. Not surprisingly, we see almost perfect convergence at these radii. In fact, because of more accurate integration and finite gravitational softening (which reduces the N-body relaxation time below the equation for strict point masses assumed by the argument in Power et al. 2003), if we are willing to tolerate slightly larger errors, within a factor ~ 1.3 (0.1 dex) of the converged (NFW) solution we find good convergence down to radii containing just $N \gtrsim 200$ particles.

Fig. 12 compares the DM-only result to the full baryonic physics runs from Fig. 8. Following standard practice, we correct the profiles from baryonic runs by the cosmic mean baryon fraction, so that if the baryons behaved identically to the DM, the curves would lie exactly on top of one another. With or without baryons, in both our dwarf and MW-mass runs, convergence is

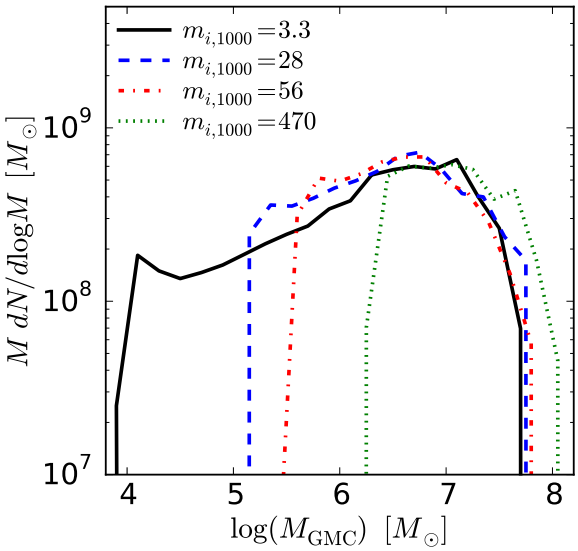

Figure 13. Resolution study of ISM structure. We show the resolution study as Fig. 10, re-running the MW-mass simulation **m12i** from $z=0.07-0$, and measure the time-averaged mass function (MF) of "GMCs": dense, cold gas clouds (found in post-processing with a friends-of-friends algorithm; see text). We plot the total ISM mass in clouds in each logarithmic interval in cloud mass ($M_{\rm GMC}\,dN_{\rm GMC}/d\log_{10}(M_{\rm GMC})$) as a function of cloud mass ($M_{\rm GMC}$). All identified "structures" as small as 3 gas particles are shown. The mass function has a power-law slope of $\sim dN/dM \propto M^{-(1.6-1.8)}$ and an exponential cutoff around the Toomre mass $\sim 10^7\,M_\odot$, very similar to both observations of MW-like galaxies and analytic predictions from turbulent fragmentation (see Blitz & Rosolowsky 2005; Murray 2011; Hopkins 2012a). Down to remarkably few particles per cloud ($M_{\rm GMC}\sim 3-5\,m_i$), the MFs appear converged; going to higher resolution simply samples further down the MF. Most of the mass in all cases is around the Toomre mass, so the total mass in clouds is within $\sim 30\%$ in all runs shown. Only the lowest-resolution run ($m_{i,\,1000}\sim 470$) cannot fully-resolve this peak, so some of the mass in clouds (which is similar) is numerically "shifted" into larger-mass structures, biasing the mean $M_{\rm GMC}$ by a factor ~ 2.

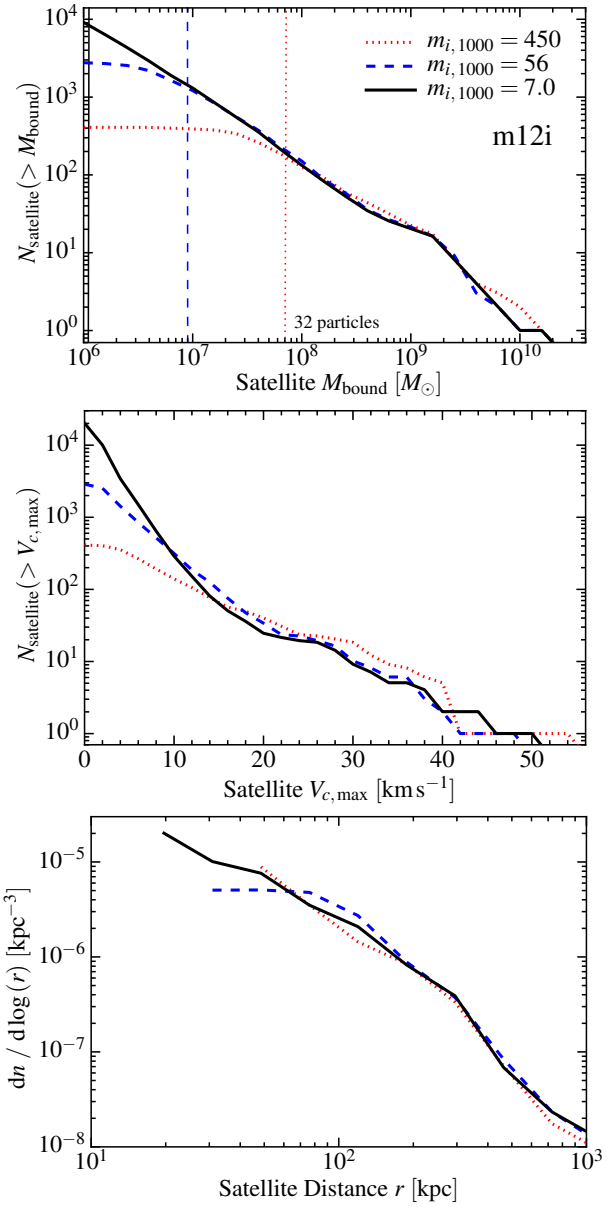

Figure 14. Mass function (*top*), circular velocity distribution (*middle*), and spatial distribution (*bottom*) of satellites (DM subhalos) within the primary halo in DM-only runs of our **m12i** simulation (at $z=0$), at varied mass resolution (as labeled). *Top:* Number of subhalos versus bound subhalo mass. Vertical lines show 32 DM particles; convergence is good down to structures containing $\sim 10-30$ particles. *Middle:* Number of subhalos versus maximum circular velocity. Convergence is similar to the mass function (deviations occur only below $\sim 10-15\,{\rm km\,s^{-1}}$, corresponding to the un-resolved subhalos above). *Bottom:* Distribution of subhalos in radial distance from the center of the primary halo. Up to shot noise in the exact position of the individual subhalos in their orbits, the distributions agree well. As the DM is approximately self-similar, we find qualitatively identical results comparing our other MW-mass (**m12f**) or dwarf mass (**m10q**, **m10v**) simulations.

good to $\lesssim 0.15$ dex outside the radii enclosing ~ 200 particles. In **m10v**, the $z=0$ galaxy is sufficiently low-mass that baryons have a negligible effect on the DM profile (see Chan et al. 2015), so the two sets of runs track each other closely. In **m10q**, the galaxy starts to reach masses where stellar feedback generates a "core": even though small (factor <2) differences in the stellar mass have large effects on core creation at these masses (see Oñorbe et al. 2015), the convergence is similar in baryonic and DM runs. In the MW-mass runs, we see that runs with baryons have higher central densities than DM-only runs, because the large galaxy baryonic mass has caused some halo contraction. This leads to nominally better convergence in an N-body sense, but again means that the DM profile is sensitive to changes in the galaxy stellar mass (here, the highest-resolution run has a lower M_* and correspondingly lower central DM density). In all cases, our highest-resolution baryonic simulations reach an approximate converged resolution of $\sim 50-100$ pc *in the dark matter*. We of course achieve much higher effective spatial resolution in the baryons dense enough to form stars.

Note that there is no analogous Power et al. (2003) convergence criterion for baryons. However, Fig. 13 and various self-gravitating baryonic collapse tests in Hopkins (2015) show that self-gravitating structure can be captured across just a couple inter-particle separations. Larger numbers of particles are needed for convergence in the mass profile in dark matter because the N-body orbits must be integrated for a Hubble time. In the gas, however, self-gravitating sub-structures (e.g. GMCs) survive for only a few dynamical times. Thus the appropriate Power et al. (2003)-equivalent criterion is $t_{\rm relax}\gtrsim$ couple $\times\,t_{\rm dyn}$, which is easily satisfied even for point masses (un-softened gravity) whenever an object is comparable in mass (or larger) than a kernel. With adaptive softening for gas (our default choice), this is *by definition* true even for structures of ~ 2 particles.

Fig. 13 specifically considers the convergence of structure *within* the ISM; we use the same simulations from Fig. 10 of the

MW-mass system at $z \sim 0$, but study the mass function of dense, cold gas structures within the ISM.[12] This is a simple proxy for e.g. the GMC mass function. The shape of the mass function is, as expected, a power-law with slope $dN/dM \propto M^{-(1.6-1.8)}$ (very similar to observed; see Blitz & Rosolowsky 2005), and turnover at a maximum mass about the Toomre mass $M_{\rm Toomre}$ defined above (also as expected from observations and analytic theory; Murray 2011; Hopkins 2012a). The lower limit is purely a resolution effect: we only keep structures with ≥ 3 particles. Remarkably, the mass function appears converged down to clouds with just a few gas particles – we simply sample further and further down the mass function as we increase the mass resolution. Since most of the mass is in Toomre-mass structures, the total mass in clouds is also converged to within $\sim 30\%$ in all the runs plotted. The only case which may be biased is the lowest-resolution example ($m_{i\,1000} = 470$), where the most massive clouds appear slightly more massive: this is because the clouds with $\sim 10^6\,M_\odot$ (not quite the peak of the MF, but just below it) contribute a significant total mass but cannot be resolved (this is ~ 2 particles), so that mass is "shifted" into more massive structures numerically. But even in this marginal case, the gas-mass weighted mean $M_{\rm GMC}$ is only over-estimated by a factor ~ 2. In our best-resolved case, $> 75\%$ of the GMC population gas mass (and $> 75\%$ of the SFR of the entire galaxy) is contained in clouds with $> 10^6\,M_\odot$, similar to what is observed in the Milky Way (Williams & McKee 1997).

Note that Hopkins et al. (2012c), using the same feedback physics (but in non-cosmological simulations using a different code), consider a much more detailed study of simulated GMC mass functions as well as size-mass relations, linewidth-size relations, virial parameters, internal column density distribution functions, lifetimes, and star formation efficiencies. Their convergence tests (and agreement with observations) are all consistent with our simple comparison here.

This further demonstrates, as we argued above, that resolving the thermal Jeans mass in cold gas is not important to cloud-or-larger scale dynamics (it only should matter if we are trying to resolve individual proto-stars). We will further demonstrate this below, when we consider adding artificial pressure floors to the ISM, or remove the low-temperature ($\ll 10^4$ K) cooling physics: because of the strong dependence of Jeans mass on pressure, these changes in turn change the thermal Jeans mass in the cold phase gas by several orders of magnitude (factors $\sim 10^3 - 10^5$), yet they have no effect on any results we measure.

Finally, Fig. 14 compares the mass function and spatial distribution of subhalos within our $z = 0$ **m12i** halo, as a function of mass resolution. Down to a subhalo mass (and corresponding maximum rotation velocity) corresponding to $\sim 10 - 30$ DM particles, the predictions are well-converged. Given the strong dependence of galaxy baryonic mass on halo mass, and much stricter resolution criteria for galaxy baryonic properties, the DM substructure mass functions and orbital distributions are always extremely well-resolved when we consider a galaxy resolved.

[12] For simplicity, we use a friends-of-friends group finder with a linking length $= 0.2$ times the mean inter-particle separation within the galaxy, to identify substructures in the gas at temperatures below ≤ 8000 K and densities $n > 5\,{\rm cm^{-3}}$, within 0.1 virial radii of the center of the main galaxy, at ~ 30 uniformly-spaced snapshots in time between $z = 0.06 - 0$, and plot the time-averaged mass function of gas structures, $M\,dN/d\log_{10}M$ for each simulation.

4.2 Spatial Resolution

Because our simulations are Lagrangian, there is no single meaningful definition of the spatial and/or force resolution. Here we discuss different criteria, and study their importance for our conclusions.

4.2.1 Hydrodynamic Resolution

The meaning of hydrodynamic resolution in our mesh-free Godunov methods is discussed extensively with dozens of numerical examples in the methods papers (Hopkins 2015, 2016a,b; Hopkins & Raives 2016). We briefly review it here. For hydrodynamics the resolution is fully adaptive, set by the inter-particle spacing h_i – it can, in principle, become arbitrarily small. In our Lagrangian method, the mean inter-particle spacing is directly related to the gas density $n_{\rm H}$, via

$$h_i^{\rm gas} = 16\,{\rm pc}\,m_{i,\,1000}^{1/3}\left(\frac{n_{\rm H}}{10\,{\rm cm^{-3}}}\right)^{-1/3} \quad (8)$$

In Hopkins (2015); Hopkins & Raives (2016) we show that the "effective" resolution for sound waves is identical to second-order grid methods such as ATHENA, meaning one element ("particle") is equivalent to one cell, and the appropriate "resolution" should be taken to be the inter-particle spacing, as opposed to the extended "kernel search radius" (maximum distance to any interacting neighbor cell), which is more akin to the gradient stencil in grid-based codes (and depends on the detailed kernel shape). This is different from SPH methods, where properties are (by definition) "smoothed" over a kernel. In fact, because of the higher-order reconstruction, sound waves can be accurately re-constructed to scales about $\sim 1/2$ the inter-particle separation (Hopkins 2015 Fig. 2, Hopkins & Raives 2016 Fig. 1). Contact discontinuities can be captured and conserved across $\sim 2-3$ elements (two inter-element spacings), much sharper than non-Lagrangian codes if the discontinuity is moving (see Hopkins 2015 Figs. 16-23). Strong shocks are resolved over a similar ~ 3 elements, close to ATHENA (Hopkins 2015 Figs. 3, 10-15, 29-30). Intrinsically multi-dimensional problems, such as conserving vorticity of a rotating, pressure-equilibrium disk for a few orbits are more demanding; this requires a few hundred total elements (about ~ 10 elements "per radian" around the structure; Hopkins 2015, Fig. 4-5). However that is easily satisfied for any disk with a resolved vertical scale-height $H \gtrsim 1 - 2\,h_i$. In multi-dimensional super-sonic turbulence, numerical dissipation and noise truncates the inertial range at wavelengths approximately ~ 5 times the inter-particle spacing (comparable to AREPO and ATHENA; Hopkins 2015 Fig. 26-28).

4.2.2 Gravitational Resolution: Star Forming, Jeans, Convergence, and Other Scales

The meaning of "gravitational spatial resolution" is somewhat ambiguous. Commonly, however, this is used to refer to the force softening ϵ. As discussed in § 2.2, we solve gravitational forces for the *same* gas-mass distribution as hydrodynamic forces: this means setting the gravitational force softening for gas adaptively,

$$\epsilon_i^{\rm gas} \equiv h_i^{\rm gas} = 16\,{\rm pc}\,m_{i,\,1000}^{1/3}\left(\frac{n_{\rm H}}{10\,{\rm cm^{-3}}}\right)^{-1/3} \quad (9)$$

For clarity, we use the same definition of force softening and spatial resolution, corresponding to the inter-particle separation $h_i = \Delta x$ (defined such that $\rho_i = m_i/h_i^3$). For our default kernel, this means that the commonly quoted "Plummer equivalent" softening is $\epsilon_{\rm plummer} \approx (2/3)\,\epsilon = (2/3)\,h_i$. For the reasons in § 2.2 and § 4.2.3 below, we adopt constant force-softening parameters for DM ($\epsilon_{\rm DM}$)

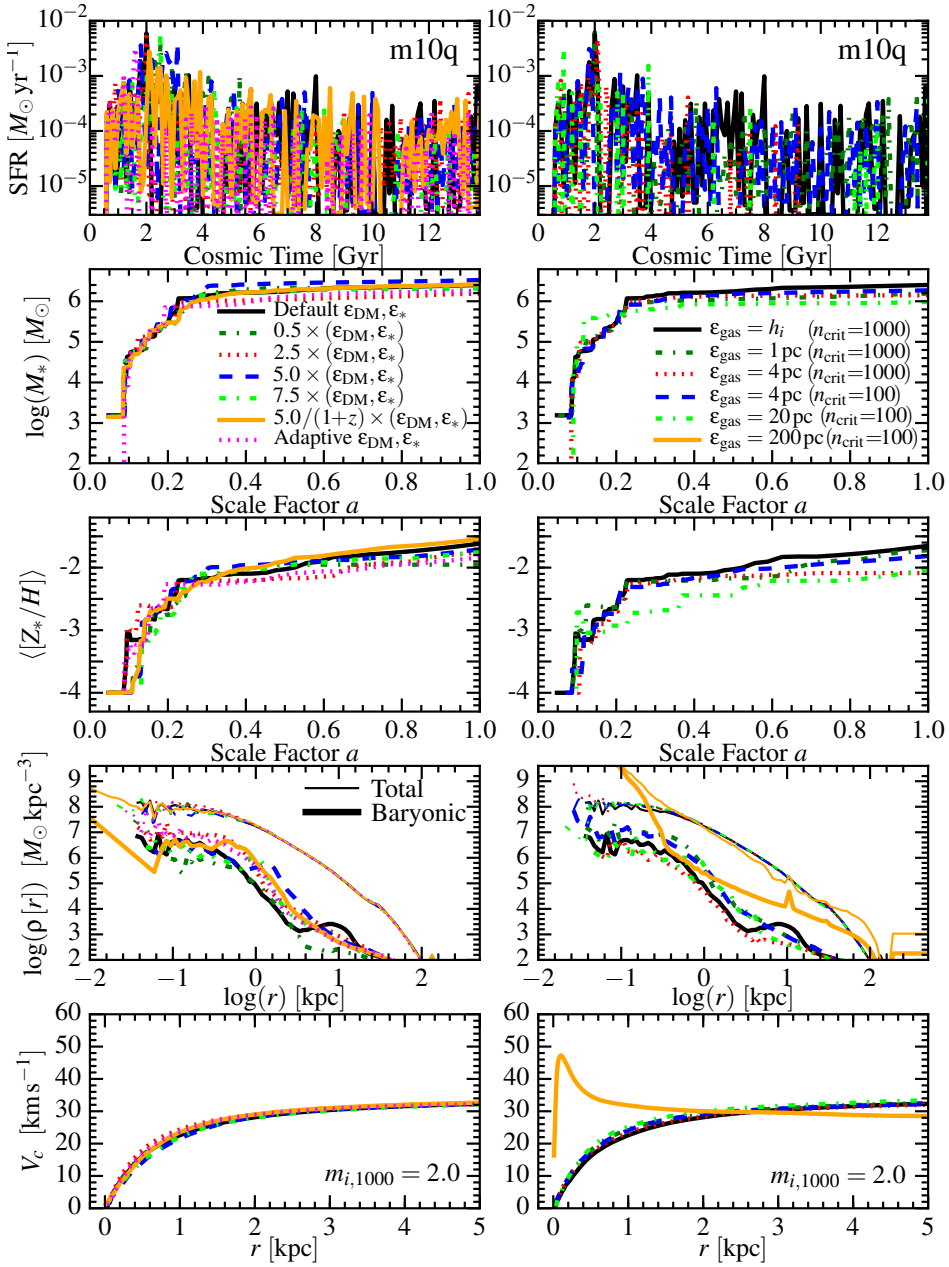

Figure 15. Spatial resolution study, as Fig. 8. We take the **m10q** simulation with fixed mass resolution $m_{i,1000}=2$, and re-run with different gravitational force softening. *Left:* Varying force softening for collisionless particles (DM ϵ_{DM} and stars ϵ_*). We multiply the default values from Fig. 8 at this mass resolution (constant $\epsilon_{DM}=40$ pc, $\epsilon_*=10$ pc) by constant values from $\sim 0.5-7.5$; we also compare softening fixed in co-moving units (default values $\times 5/(1+z)$); and we consider fully adaptive ϵ_{DM} and ϵ_*. These changes have no systematic effect (small differences in mass are consistent with stochastic variations, given the bursty SFH). We have verified the same in more limited surveys of ϵ_{DM} and ϵ_* in simulations **m10y**, **m11q**, **m11v**, **m12i**. *Right:* Varying force softening for gas. We compare our default, fully-adaptive softening ($\epsilon_{gas}=h_i$, the inter-particle spacing), to simulations with fixed gas softening set to $1-200$ pc. Our default (adaptive) run and those with $\epsilon_{gas}=1,4$ pc enforce a minimum gas density for star formation $n_{crit}=1000\,cm^{-3}$ (as labeled). With 1 pc softening, this produces identical results to our adaptive-softening run; but with 4 pc softening, the maximum gravitationally-resolved density $n^{max}=(m_i/\epsilon_{gas}^{min})/m_p\approx 1000\,cm^{-3}\,(m_{i,1000}/2)\,(\epsilon_{gas}^{min}/4\,pc)^{-3}$ barely reaches n_{crit}, so our self-gravity criterion cannot properly resolve if gas is self-gravitating at $n>n_{crit}$ and SF is artificially suppressed (slightly lower M_* and $[Z_*/H]$). Lowering $\epsilon_{gas}=20$ pc ($n^{max}\approx 10\,cm^{-3}$) with $n_{crit}=100$ shows this problem more severely. With fixed (non-adaptive) softening, it is necessary to choose $n_{crit}<n^{max}$; lowering $n_{crit}=100$ for $\epsilon_{gas}=4$ pc shows excellent agreement with our default (adaptive) run. With $\epsilon_{gas}=200$ pc, $n^{max}=0.01\,cm^{-3}$ is much less than n_{crit} *and* the mean galaxy gas density – this means we cannot resolve Toomre-scale structures (large GMCs), so star formation is completely suppressed ($M_*=0$, exactly, here), and a super-dense (but not locally-fragmenting) gas disk forms.

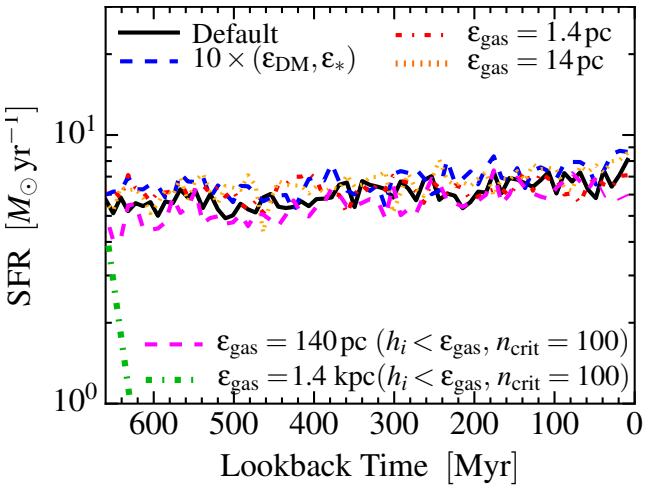

Figure 16. Spatial resolution study in re-starts of a MW-mass galaxy (**m12i**) at late times, as Fig. 10, keeping fixed mass resolution $m_{i,\,1000}=56$ but varying the spatial resolution, as in Fig. 15. All runs unless otherwise labeled take $n_{\rm crit}=1000\,{\rm cm}^{-3}$. Our default run uses adaptive softening for gas, fixed softening for DM and stars ($\epsilon_{\rm DM}$, ϵ_*); the run labeled $10\times(\epsilon_{\rm DM},\,\epsilon_*)$ increases these by $10\times$, but keeps the adaptive gas softening – this has no effect on our prediction. Runs labeled $\epsilon_{\rm gas}$ fix the gas softening; for $\epsilon_{\rm gas}=1.4\,$pc and $14\,$pc ($n^{\rm max}=10^6, 10^3\,{\rm cm}^{-3}$, respectively), this has little or no effect. For $\epsilon_{\rm gas}=140\,$pc ($n^{\rm max}\approx 1\,{\rm cm}^{-3}$), we see no star formation unless we lower $n_{\rm crit}$ and allow the hydrodynamic resolution h_i to decrease to values much smaller than $\epsilon_{\rm gas}$ ($<0.1\,\epsilon_{\rm gas}$, here) – then shocks driven by SNe produce gas at high densities (i.e. hydrodynamic, not gravitational, effects allow the gas to reach high densities), but the most massive GMCs are still only marginally resolved. For $\epsilon_{\rm gas}\sim$kpc, the galaxy never forms locally self-gravitating structures, even lowering $n_{\rm crit}=1$ and allowing $h_i=0.01\,\epsilon_{\rm gas}$.

and stars (ϵ_*), with their values set to the "optimal" values recommended below (for specific values, see Table 1). We define these for clarity the same way as for gas, so that they are softened according to the same kernel shape (in other words, a gas, star or DM particle with the same mass and ϵ will exert identical gravitational forces at all radii); this means once more that the equivalent Plummer softening is $\approx(2/3)\,\epsilon$.

Below, we will vary these parameters and show that, provided reasonable values are chosen (sufficiently small to avoid over-softening real features in the system, such as resolved GMCs), the force softenings ϵ play a relatively minor role in convergence of our results. And for adaptive softenings (e.g. gas), there is no single "spatial resolution" – the smallest values allowed may never be reached by the simulation, and the smallest values reached may only be reached by a tiny particle number. Moreover, § 4.1.4 shows that mass resolution may be what actually determines the effective spatial resolution of DM mass profiles (and internal kinematics, in our tests).

In Table 3, we therefore provide a number of other "effective spatial resolution" values for one of our resolution studies (our **m12i** study) from Fig. 8, at each resolution level we have considered. Specifically we quote the "spatial resolution" (inter-particle spacing h_i) at:

(i) **Maximum Density/Minimum Softening:** This is simply the minimum $\epsilon_i^{\rm gas}\equiv h_i^{\rm gas}$ reached in the simulation, at any time. As noted above, this (by definition) represents the extreme in the simulation, not "typical" values – in our highest-resolution simulations, this corresponds to densities $n_{\rm max}\sim 5\times10^6\,{\rm cm}^{-3}$, which are

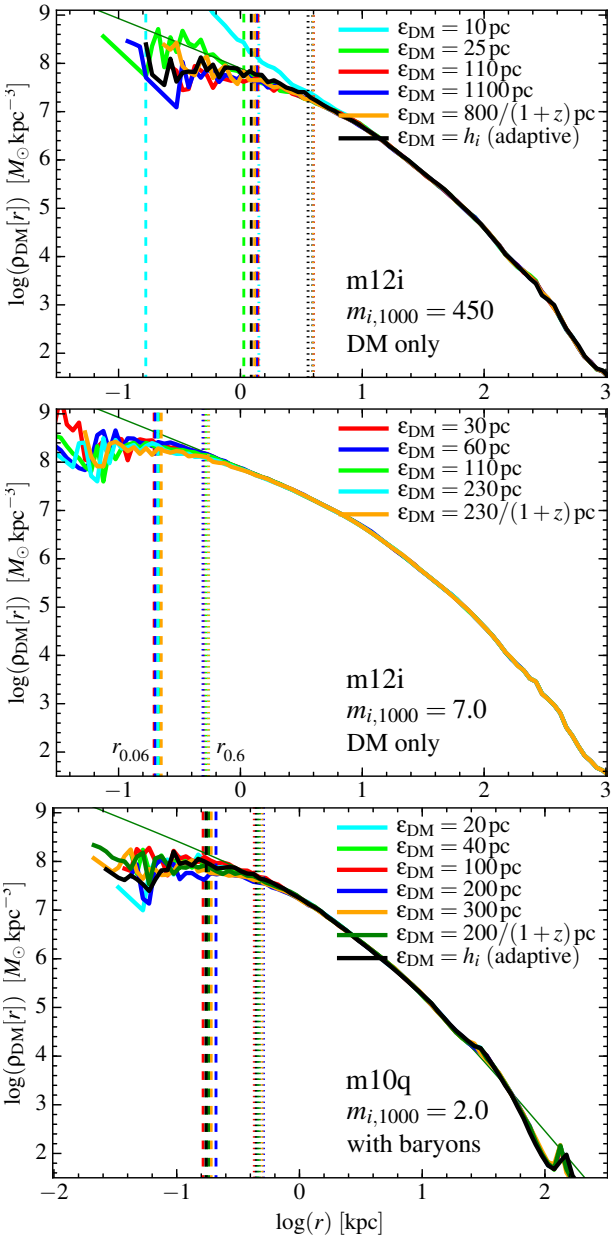

Figure 17. Effects of force softening on the DM profile at $z=0$, at fixed *mass* resolution. *Top:* DM-only simulation of **m12i**, with $m_{i,\,1000}=450$. We systematically vary the DM force softening and consider a co-moving and fully-adaptive case. *Middle:* DM-only simulation of **m12i**, with high resolution $m_{i,\,1000}=7$. *Bottom:* Full-physics simulation of **m10q**, with intermediate resolution ($m_{i,\,1000}=2$). Vertical dashed (dotted) lines show the convergence radii $r_{0.06}$ ($r_{0.6}$), as Fig. 11; thin green line is the best-fit NFW profile to the highest-resolution simulation we have run. The choice of DM softening with $\epsilon_{\rm DM}\sim20-1000\,$pc ($\sim0.05-5$ times the mean inter-particle spacing in the central $\sim1\,$kpc) has no measurable effect within $r_{0.6}$ and only small ($<0.2\,$dex) effects within $r_{0.06}$. At each resolution, the largest force softenings shown produce a small suppression of the central density ($\sim0.2\,$dex between $r_{0.06}$ and $r_{0.6}$) relative to the converged solution; this occurs when $\epsilon_{\rm DM}\gtrsim0.5\,r_{0.06}$. Once $\epsilon_{\rm DM}\lesssim0.25\,r_{0.06}$, we see no further improvements in convergence. Only for the smallest $\epsilon_{\rm DM}$ at our lowest resolution ($\epsilon_{\rm DM}=10\,$pc, $m_{i,\,1000}=450$, **m12i**) do we see hard-scattering producing numerical N-body relaxation (a spurious cusp) – this case has $\epsilon_{\rm DM}<0.01\,r_{0.06}$ (for $r_{0.06}$ of well-behaved solutions), and <0.05 times the inter-particle spacing.

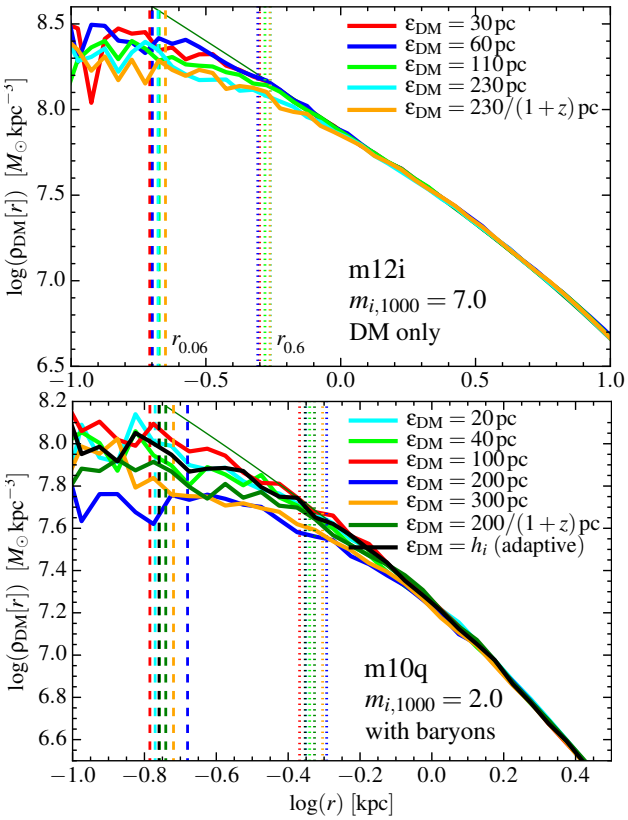

Figure 18. Close-up of the two higher-resolution simulations from Fig. 17, showing the DM density profile in the halo center in more detail. We can now clearly see that while the profiles agree nearly perfectly at large radii ($\gtrsim$ kpc), at small sub-kpc radii, the runs with excessively large force softening $\epsilon_{\rm DM}$ suppress the density profile relative to the converged and higher-resolution solutions, by a modest factor ~ 1.2 (~ 1.5) at $\sim 500\,$pc ($\sim 100\,$pc). The suppression can even extend to radii of order $r_{0.6}$ (*dotted vertical line*). The excess suppression appears in all cases we have tested when $\epsilon_{\rm DM} \gtrsim 0.5\,r_{0.06}$ (our convergence radius $r_{0.06}$ is shown as the vertical dashed line) – approximately $\epsilon_{\rm DM} \gtrsim 100\,$pc in both cases shown here. Once $\epsilon_{\rm DM}$ is less than $\sim 0.25\,r_{0.06}$, we see no evidence for further convergence in the simulations (differences between these runs and adaptive-softening runs are consistent with shot noise).

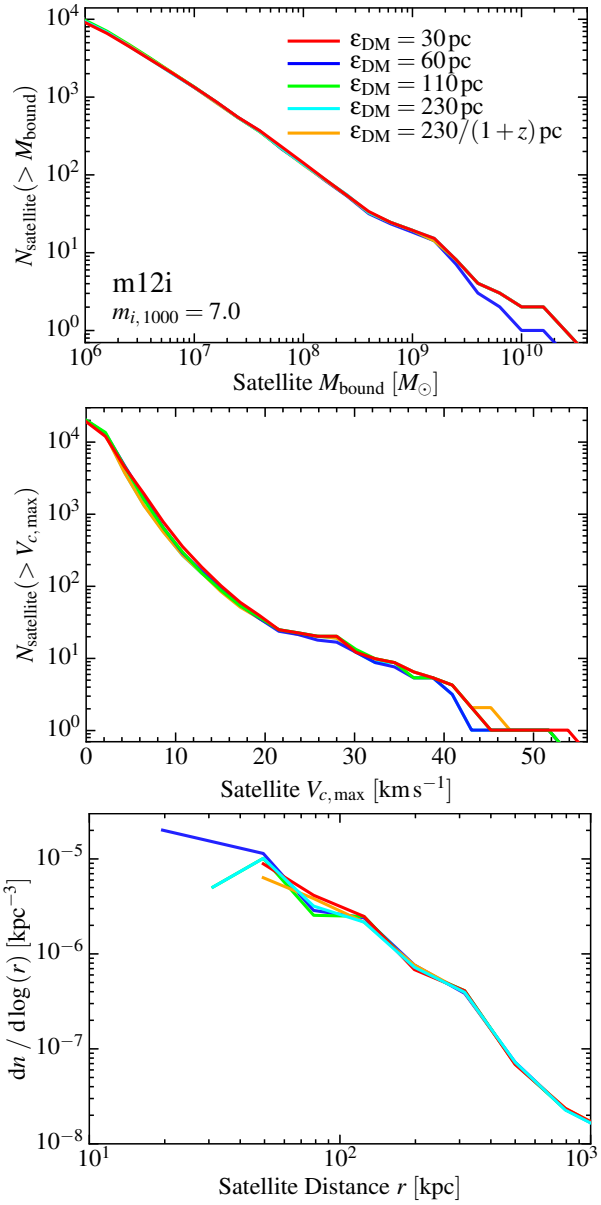

Figure 19. Mass function (*top*), circular velocity distribution (*middle*), and spatial distribution (*bottom*) of satellites (DM subhalos) within the primary halo in DM-only runs of our **m12i** simulation (at $z = 0$), as Fig. 14, at fixed mass resolution ($m_{i,1000} = 7$) but varied force softening $\epsilon_{\rm DM}$. *Top:* Number of subhalos versus bound subhalo mass. Objects are plotted down to < 30 DM particles. *Middle:* Number of subhalos versus maximum circular velocity. *Bottom:* Distribution of subhalos in radial distance from the center of the primary halo. All the distributions agree nearly perfectly, independent of force resolution (the small deviation in the mass function of $\epsilon_{\rm DM} = 60\,$pc owes to a single subhalo which falls just outside, instead of inside, the radius cut used to identify substructures at $z = 0$).

reached by the simulation only briefly in the galactic nucleus in an intense, high-redshift starburst.

(ii) **Star-Forming Densities:** We output from our simulations the gas density at which every individual star particle forms, which maps one-to-one to the equivalent h_i. We show the mean density, $\langle n_{\rm SF}\rangle$, weighted by total star formation, integrated over the entire galaxy history to $z = 0$, and corresponding $h_i^{\langle SF\rangle}$. We emphasize that because star formation is only allowed in self-gravitating gas which is also self-shielding and molecular, the mean $\langle n_{\rm SF}\rangle$ is always significantly larger than the *minimum* density at which star formation is allowed ($n_{\rm SF,\,min}$), even when the latter is set to $n_{\rm SF,\,min} = 1000\,{\rm cm^{-3}}$; for completeness we quote the equivalent $h_i^{\rm threshold}$ at this minimum density as well, but note that most stars form at higher densities in resolved, self-gravitating structures. For our highest-resolution MW-mass galaxy, most star formation occurs in regions with spatial resolution $< 5\,$pc.

(iii) **Thermal Jeans Mass (Warm/Hot Gas):** As we discuss in § 4.1.3, resolving the thermal Jeans mass is not dynamically important in galaxy simulations – the smallest resolveable Jeans mass simply tells us where the "fragmentation cascade" in self-gravitating objects will be truncated (and our sub-grid star formation model will take over). We can calculate this minimum scale, but the correct "effective Jeans mass" (the actual characteristic mass of structures) depends on turbulence, *not* on the sound speed, in a super-sonically turbulent medium (see e.g. Padoan & Nordlund 2002; Hennebelle & Chabrier 2008; Hopkins 2012b,a, 2013d,c,b; Hopkins & Christiansen 2013; Guszejnov & Hopkins 2015, 2016). To get some insight, however, for the warm ($\gtrsim 10^4\,$K) gas, the turbulence is expected to be trans-sonic, so we can reasonably sim-

plify by considering the thermal Jeans properties. We calculate the smallest possible Jeans mass resolvable with at least ~ 10 elements at these temperatures, and the corresponding resolved Jeans length (with $\sim 3\,h_i$) and gas density.[13] Any *lower* density or *larger* Jeans lengths have resolved thermal Jeans fragmentation – important for our purposes, these values need to be sufficient to capture the largest GMCs (sizes $\sim 100-200\,$pc and densities $\sim 10\,\mathrm{cm}^{-3}$), which form out of the warm gas. Note that if the gas is hotter, it becomes *easier* to resolve Jeans-scale structures.

(iv) **Turbulent Jeans Mass (Cold Gas):** Once the gas cools below temperatures $\sim 10^4\,$K (in e.g. GMCs or molecular disks), the turbulence is highly super-sonic – by definition, thermal pressure does not control the dynamics, and the appropriate Jeans scale for fragmentation is the *turbulent* Jeans scale. This replaces $\lambda_J = c_s/\sqrt{G\rho}$ by $\lambda_{\rm turb} \approx \langle v_{\rm turb}^2(\lambda_{\rm turb})\rangle^{1/2}/\sqrt{G\rho}$, where $v_{\rm turb}(\lambda_{\rm turb})$ is the rms turbulent velocity measured within the same region. A rigorous definition and derivation of the corresponding turbulent Jeans length, mass, and fragmentation cascade is given in Hopkins (2013b). If we assume a linewidth-size relation seen in the ISM and expected for super-sonic turbulence ($v_{\rm turb} \propto \lambda^{1/2}$), and that the "parent" clouds have virial parameter of unity and (according to Larson's Laws) a universal surface density $\Sigma_{\rm cloud} = \Sigma_{300}\,300\,M_\odot\,{\rm pc}^{-2}$, and take a minimum resolved-object mass $\sim N\,m_i$, then following Hopkins (2013b) gives a minimum-resolved length $\lambda_{\rm turb} \approx 4.0\,{\rm pc}\,(m_{i,\,1000}\,N_{10}/\Sigma_{300})^{1/2}$ and maximum-resolved density $n_{\rm max}^{\rm turb} \approx 1.2\times10^4\,\Sigma_{300}\,(m_{i,\,1000}\,N_{10}/\Sigma_{300})^{-1/2}$.

(v) **Dark Matter Power-Type (Convergence) Radius:** None of these criteria apply to dark matter. We therefore quote the DM interparticle separation: both its minimum value $h_i^{\rm DM,\,min}$, as well as rms value within twice the effective radius of the baryonic galaxy at $z=0$, $h_i^{\rm DM,\,core}$. However for the DM profiles, § 4.1.4 above shows that internal structure of collisionless objects evolved for a Hubble time is well-converged inside a radius enclosing ~ 200 DM particles. We therefore quote this as well, for the DM.

Qualitatively, the results from our other MW-mass series (**m12f**) are nearly identical to those in Table 3. Our dwarfs (**m10v** and **m10q**) have superior resolution, in every category, owing to their much smaller particle masses.

4.2.3 "Optimal" Force Softening Choices

Although we will show that mass resolution is what primarily determines the "effective spatial resolution" of our simulations, it is still, of course, necessary to define some force resolution parameters, particularly for collisionless particles which use a constant force-softening. We will test these in detail below, but summarize our best estimate of "optimal" parameters here:

- **Gas:** For gas, the optimal force softening is un-ambiguous: it should be set adaptively to match the hydrodynamic resolution. This is optimal for several reasons. (1) It is unquestionably the physically correct set of equations for a collisional fluid. The hydrodynamic solver describes a mass distribution (not just at particle locations, but everywhere in the domain, according to the reconstruction of the method). If an element of mass m_i has hydrodynamic resolution scale h_i, *the mass is contained within that volume*, and as that element is moved, compresses, or expands, *so does the gas mass*. We should therefore solve the Poisson equation *for the same mass distribution*. It is fundamentally ill-posed to solve the Poisson equation for a different mass distribution, arbitrarily chosen to be different from the hydrodynamics.[14] (2) Many authors have shown that matching these resolution elements provides an optimal numerical softening, in the sense that it automatically provides the most accurate solution (converges most rapidly) while minimizing N-body integration errors (Merritt 1996; Bate & Burkert 1997; Romeo 1998; Athanassoula et al. 2000; Dehnen 2001; Rodionov & Sotnikova 2005; Price & Monaghan 2007; Barnes 2012; Hubber et al. 2013). (3) It guarantees that numerical hard gas-gas scattering can *never* dominate over physical self-gravitating motions. (4) It removes ambiguity about the meaning of force/spatial/hydrodynamic resolution (it makes the mass resolution the un-ambiguous resolution scale). (5) In Lagrangian codes, spatially, self-gravitating objects will be able to collapse (correctly) to arbitrarily high densities, where some star formation criterion should identify them. This makes it essentially impossible to artificially suppress star formation by "tuning" some star formation threshold (which is possible if gravity is artificially softened on large scales, preventing small self-gravitating structures from collapsing). (6) It automatically behaves correctly in all density regimes, and naturally removes any ambiguity about co-moving or physical softenings in cosmological integrations.[15]

  Obviously, if constant gravitational softenings for gas must be used (though we do not see any possible circumstances where they would), they must be chosen sufficiently small to resolve the vertical scale-heights of the cold gas disk, and Toomre lengths of the most massive Toomre-mass objects discussed in our mass-resolution criteria ($\epsilon_{\rm gas} \ll 100\,$pc).

- **Dark Matter:** For dark matter, we will show that the force softening makes little or no difference to any conclusions, so long as it is not extremely small (about two orders of magnitude smaller than any FIRE-1 or FIRE-2 simulation) which would trigger hard-scattering effects, or extremely large (which would over-soften the central DM cusp of the galaxy). One option is to use fully adaptive softening; however we will show this requires stronger timestep constraints be imposed on the collisionless fluid (Courant-like conditions which prevent the softening from changing too much within a timestep), in order to converge properly, which considerably increases the computational expense for no apparent improvements in accuracy. Moreover, for a collisionless fluid, the "correct" adaptive scaling is *physically* ambiguous. It is not simply the case that the spatial domain of the DM represented by particle a must shrink, if the DM particles b around a move inwards (i.e. if $\nabla\cdot\mathbf{v}_{\rm DM}^a < 0$),

[13] Define the Jeans length λ^J and mass $m_J = (4\pi/3)\,\rho\,(\lambda^J/2)^3$. Following § 4.1.1, a "resolved" structure has some number $N \sim 10N_{10}$ elements, so mass $\sim N m_i$. By definition of the inter-particle spacing, we also have $\rho = m_i/h_i^3$. This gives $\lambda_J \approx 3\,h_i\,N_{10}^{1/3}$, always – in other words, our adaptive softening ensures the Jeans length λ_J is always spatially-resolved *provided* the Jeans mass m_J is mass-resolved. Now use this and the definition of Jeans length, $\lambda^J = c_s/\sqrt{G\rho}$, to solve for the Jeans radius $\lambda^J/2 \approx 9\,G\,m_i\,N_{10}/c_s^2$; for $T = 10^4\,K$ (and assuming fully-ionized gas), this becomes $\approx 0.3\,{\rm pc}\,m_{i,\,1000}\,N_{10}$. Taking a minimum resolved size $N_{10}=1$, this gives the minimum resolved λ_J; the corresponding h_i gives, in turn, the corresponding maximum density $n_{\rm max}^J$.

[14] Strictly speaking, it is only possible to solve the Poisson equation for a mass distribution accurate to the same order as the reconstruction order of the hydrodynamic method. But up to truncation errors, this is exact in matching the gas distribution.

[15] Adaptive gas softening is also, in fact, already the standard in most grid-based codes (Kravtsov et al. 1997; Teyssier 2002; Bryan et al. 2014) and moving-mesh codes (Springel 2010) and in particle-based codes as well in e.g. the fields of star and planet formation (see e.g. Bate & Burkert 1997). Particle-based codes in galaxy formation have proved a surprising historical exception.

because those particles can move *through* particle a (they do not, in fact, have to compress it, in configuration/position space).

Therefore, fixed softenings appear preferable. We adopt softenings fixed in physical units at $z < 10$, since the halo centers do not change significantly in density over this redshift range, so this is closer to the typical softenings returned by the adaptive-softening method (i.e. closer to a fixed ratio relative to the inter-particle separation; co-moving softenings are only relevant far out in the IGM in the linear regime, outside our high-resolution box entirely).

For fixed DM softenings, avoiding over-softening in the central DM profile – i.e. converging as accurately as possible to the solution of higher-resolution simulations – requires a force softening *smaller* than the 200-particle Power-like radius (where $t_{\rm relax} \approx 0.06\,t_{\rm circ}(R_{200})$) by a factor of at least a couple, i.e. $\epsilon_{\rm DM} \lesssim 0.5\,r_{0.06}$. This is approximately ensured if $\epsilon_{\rm DM} < 30\,{\rm pc}\,m_{i,1000}^{1/2}\,(M_{\rm vir}/10^{12}\,M_\odot)^{-0.2}$ (we estimate this by taking $M_{\rm enc}(<r_{0.06}) \approx 220\,m_i$, and assuming NFW profiles with concentration $c \approx 10\,(M_{\rm vir}/10^{12}\,M_\odot)^{-0.15}$). The minimum softening is given by the value where we begin to see hard scattering effects cause excessive N-body relaxation in the profile and/or velocity distribution function, which is roughly $\epsilon_{\rm DM} > 0.03\,{\rm pc}\,m_{i,1000}\,(M_{\rm vir}/10^{12}\,M_\odot)^{-2/3}$ (obtained by comparing the hard-scatter Δv to V_c). Note the extremely large dynamic range in between: as we will show, we can vary $\epsilon_{\rm DM}$ by multiple orders of magnitude and obtain identical converged results. For all of our high-resolution runs, any softening in the range $1\,{\rm pc} \lesssim \epsilon_{\rm DM} \lesssim 75\,{\rm pc}$ satisfies these.[16] To be conservative, and avoid excessively small timesteps, in our fixed-softening runs we choose softenings about $\sim 2-4$ times smaller than the upper limit above (the largest softening that does not over-soften the Power-type radii), corresponding to the inter-particle (DM) spacings at the halo center (for a typical central DM density of $\sim 10^8\,M_\odot\,{\rm kpc}^{-3}$, this is $\epsilon_{\rm DM} \sim 20\,{\rm pc}\,m_{i,1000}^{1/3}$). Various tests have shown this is optimal, in fixed-softening situations to reduce both noise and over-softening errors (see, e.g. Merritt 1996; Athanassoula et al. 2000; Dehnen 2001; Barnes 2012; Hubber et al. 2013).[17]

- **Stars:** Stars are the most ambiguous softening case. Fortunately, like dark matter, our results appear almost completely independent of how gravity from stars is softened (we will show both adaptive softening and factor ~ 30 changes in fixed ϵ_*, similar to our DM studies above). As with DM, the same physical ambiguities apply to using adaptive softenings for a collisionless fluid. Moreover, if the adaptive softening is based on the inter-star distance, a star born in a pure gas cloud would instantly "jump" to a huge force softening; if based on inter-gas distance, a pure-stellar bulge would be over-softened out to the gaseous halo. In runs with adaptive DM+stellar softening, we therefore set ϵ_* to scale with the inter-baryon distance (gas+stars), which at least has the advantage that it handles both extremes above correctly. Even in this case, however, a "real" (self-gravitating, resolved) star cluster can artificially expand once gas is blown away, because the neighbor search expands, even though self-gravity should fix the size (hence softenings). With constant softenings, the ambiguity is that stars form with very different densities. We have also tested $\epsilon_*(i)$ fixed in time but variable particle-to-particle, to equal the softening of the gas from which the particle formed; while this does not change our results, it leads to wildly variable ϵ_* for stars in the same location at late times (which formed from gas at different densities at different times), which is numerically problematic. A fixed ϵ_* chosen to be large ensures smoothness of the potential but this is not actually physical for stars (because star formation is clustered, the stellar potential is "lumpy" on small scales), and we see this again causes sudden, massive expansion of the softening after a star particle forms. A reasonable compromise is to set ϵ_* similar to the gas softening at the mean density of star formation, $\epsilon_* \sim 3.4\,{\rm pc}\,m_{i,1000}^{1/3}\,(\langle n\rangle_{\rm SF}/10^3\,{\rm cm}^{-3})^{-1/3}$, and our default choice follows this criterion.

4.2.4 *Convergence Tests, N-Body Effects, and Effects of Collisionless Softening*

Figs. 15-16 consider the effects of varied force softening for both collisional (gas) and collisionless (DM, stars) particles in our cosmological simulations, for both a dwarf and MW-mass galaxy. First, consider the effects of the collisionless softening. We vary both DM and stellar softening simultaneously by multiplying both by a constant ($\sim 0.5-10$) relative to their default values, setting them to be fixed in comoving (instead of physical) units at all redshifts, or setting them purely adaptively (with appropriately careful time-stepping, as described below).

To first approximation, we see no effect from these changes. There are some variations in the growth history in Fig. 15, but given the extremely bursty nature of dwarf star formation histories, this appears to be primarily stochastic. Still, there is a (weak) tendency, on average, for the runs with larger softenings to produce slightly higher stellar masses (and corresponding metallicities). Although not apparent by-eye in Fig. 15, these larger-softening runs appear to have slightly less-bursty SFHs (if we quantify this by measuring the logarithmic variance in the SFR measured in 10 Myr bins relative to a rolling Gyr-average value) – and in previous studies, we have shown that more-bursty SF produces more efficient outflows (because the same feedback is more concentrated; Muratov et al. 2015). We have also verified this by re-running our **m11q** and **m11v** simulations (the runs where "burstiness" has the most dramatic effect on the DM halo structure) with $5\times$ larger softening $(\epsilon_{\rm DM}, \epsilon_*) = (200, 20)\,{\rm pc}$; the "burstiness" is still obviously present but is quantitatively suppressed, and the stellar masses of both increase by a factor ~ 1.4. It makes sense that such large softenings suppress burstiness to some extent: they smear out the DM mass profile, meaning the potential is weaker (so SF needs to be less concentrated/dramatic before it can drive outflows) and smear out any star clusters or small merging galaxy stellar components that would otherwise provide more concentrated feedback. But these

[16] Note that it becomes impossible to satisfy both criteria for particle masses m_i corresponding to $\lesssim 40$ DM particles in the halo; of course this is because the halos then cannot be internally spatially resolved.

[17] Another common criteria for DM softening is that the maximal 2-body acceleration ($a_i \sim G m_i/\epsilon_{\rm DM}^2$) not exceed the bulk acceleration at some radius ($a_{\rm bulk} \sim G M_{\rm enc}(<r)/r^2$) in the halo. If we assume an NFW profile, and require $a_i < \langle a_{\rm bulk}\rangle$ where $\langle a_{\rm bulk}\rangle$ is the mass-averaged mean acceleration of DM within the halo, we obtain $\epsilon_{\rm DM} \gtrsim 6\,{\rm pc}\,(c/10)^{-0.45}\,m_{i,1000}^{1/2}\,(M_{\rm vir}/10^{12}\,M_\odot)^{-1/2}$ (we approximate the exact scaling with the halo concentration c by a power-law, good to $\sim 10\%$ over the entire range of interest). Our fixed-$\epsilon_{\rm DM}$ choices agree reasonably well with this scaling. However, we stress that we see no measurable errors or deviations in the mass profile, rate-of-growth of structure, or velocity distribution function, even at/outside $R_{\rm vir}$, using order-of-magnitude smaller $\epsilon_{\rm DM}$. This is because this acceleration criterion is not meaningful when $\epsilon_{\rm DM}$ is below the inter-particle separation, because only very rarely will particles approach within separations $\ll \epsilon_{\rm DM}$, and when they do, the net *velocity* imparted by the encounter will be given by the scaling in the text: $(\Delta v/\sigma_{v,\rm DM}) \sim 10^{-4}\,m_{i,1000}\,(20\,{\rm pc}/\epsilon_{\rm DM})\,(\sigma_{v,\rm DM}/100\,{\rm km\,s^{-1}})^{-2} \ll 1$. Equivalently, the timescale required for N-body heating at $\sim R_{\rm vir}$ to perturb a particle orbit by $\sim 10\%$ is much greater than $t_{\rm Hubble}$ for all $\epsilon_{\rm DM} \gtrsim 1\,{\rm pc}$.

are unphysical effects; furthermore, the "level of burstiness" in the adaptive softening runs agrees well with our default simulations – together this gives us confidence in our default choices. But in any case, the magnitude of this effect is much smaller than systematic uncertainties in the stellar mass predictions. Not surprisingly, for the MW-mass system, which has a more smooth SFH, the effect is minimal ($\sim 10\%$).

Fig. 17 confirms the result of previous studies (e.g. Power et al. 2003) that at fixed mass resolution, changing the DM force softening (either in physical or comoving units or using fully-adaptive softenings) has very little effect on convergence. We see that outside the radius containing $N \sim 200$ particles, these choices (with fixed $\epsilon_{\rm DM}$ varied from $\sim 10-1000\,$pc) have almost no effect on the mass profile. Of course, if $\epsilon_{\rm DM}$ is too large, it will eventually suppress any smaller-scale structure – this occurs when $\epsilon_{\rm DM}$ is larger than the convergence radius $r_{0.06}$. Power et al. (2003) show that convergence to the correct solution in the central structure of N-body dark matter halos also requires $\epsilon_{\rm DM} < c^{-2}\,[\ln(1+c) - c/(1+c)]\,R_{200} \lesssim 0.01\,R_{200}$; this is easily satisfied by any of our simulations that satisfy $\epsilon_{\rm DM} \lesssim r_{0.06}$. Note that most previous studies of this have considered only pure-DM simulations; we show here the same conclusions apply even in our "full physics" runs.

Fig. 17 does raise one important caveat: if $\epsilon_{\rm DM} \ll h_i$ (the inter-particle spacing), and the "hard scattering" velocity deflection between two particle is comparable to their velocity dispersion, then runaway N-body effects can produce a gravito-thermal catastrophe over a Hubble time. We see this in our test with $\epsilon_{\rm DM} = 10\,$pc and $m_{\rm DM} = 2.8\times10^{6}\,M_{\odot}$; much smaller force softening for dark matter than used in our production FIRE-2 simulations at comparable mass resolution. Considering the hard-scattering velocity from our N-body heating rate calculation below, we estimate that avoiding this requires $\epsilon_{\rm DM} \gtrsim 0.02\,{\rm pc}\,m_{i,\,1000}$, easily satisfied in all our production simulations.

Fig. 18 shows the central portion of the profiles from Fig. 17 in closer detail. Here we can see that there is some systematic suppression of the predicted mass profile, relative to the high-resolution converged solution, in runs with large softenings $\epsilon_{\rm DM} \gtrsim r_{0.06}$. Although the effect is small, this is expected: if the profiles are converged down to $\sim r_{0.06}$ with appropriate softening, then over-softening by setting $\epsilon_{\rm DM} \gtrsim r_{0.06}$ will necessarily smear the profile out at these radii.

Fig. 19 shows the subhalo mass function, $V_{\rm max}$ function, and spatial distribution (as Fig. 14), at fixed mass resolution but varying again the DM force softening. These are almost perfectly insensitive to the DM force softening for reasonable choices. We have also examined halo formation times and internal kinematics, and find the same result.

4.2.5 *Effects of Gas Force Softening: The Dangers of Non-Adaptive Softenings*

As noted previously, Hopkins (2015) present a large number of tests demonstrating the high accuracy and near-ideal convergence rate of the implementation of adaptive gas softening in GIZMO on test problems with *known* solutions, including self-gravitating polytropic collapse (Evrard 1988), cosmological collapse of Zeldovich pancakes (with baryons and with/without DM), and steady-state orbit integration of stable (Toomre $Q > 1$) Keplerian disks, as well as good agreement with other state-of-the-art codes such as AREPO on popular code-comparison tests such as the adiabatic "Santa Barbara Cluster" (Frenk et al. 1999).

In Figs. 15-16, we also considered the effects of changing the gas force softening, replacing our default self-consistent adaptive softening ($\epsilon_{\rm gas} = h_i$) with a fixed, constant physical softening ($\epsilon_{\rm gas}$). In both our dwarf and MW-mass simulations, we see, reassuringly, that for sufficiently small fixed $\epsilon_{\rm gas}$, the differences are essentially negligible (entirely consistent with stochastic fluctuations).

However, recall that our star formation model is based on identifying self-gravitating gas above some density threshold. At fixed mass resolution, a fixed $\epsilon_{\rm gas}$ sets a minimum inter-particle separation, hence maximum gas density, at which self-gravity will be correctly calculated: this density is $n^{\rm max} = m_p^{-1}\,m_i\,(\epsilon_{\rm gas}^{\rm min})^{-3} \approx 1000\,{\rm cm^{-3}}\,m_{i,\,1000}\,(\epsilon_{\rm gas}^{\rm min}/3.5\,{\rm pc})^{-3}$. So not surprisingly, when we make $\epsilon_{\rm gas}$ large enough that $n^{\rm max} \lesssim n_{\rm crit}$ (the minimum density for star formation), we can artificially suppress the SFR. Obviously, one should never, with fixed $\epsilon_{\rm gas}$, choose $n_{\rm crit} \gtrsim n^{\rm max}$. If we lower $n_{\rm crit}$ sufficiently so that $n_{\rm crit} \ll n^{\rm max}$, then we recover nearly-identical behavior to our default simulations with adaptive $\epsilon_{\rm gas}$ and larger $n_{\rm crit}$.

This is important: we will show (and have repeatedly shown in previous papers) that the value of $n_{\rm crit}$ is almost completely irrelevant to our galaxy-scale predictions (§ 8). However, simulations in the literature which adopt fixed $\epsilon_{\rm gas}$ will see a completely artificial dependence of their results on $n_{\rm crit}$, if it approaches this limit ($n_{\rm crit} \gtrsim n^{\rm max}$).

For example, Shen et al. (2014) simulate dwarfs with mass similar to our **m10** run with fixed $\epsilon_{\rm gas} \approx 120\,$pc (correcting to our definition of ϵ), $m_{i,\,1000} = 3.3$, and $n_{\rm crit} = 100\,{\rm cm^{-3}}$ (they do not use a self-gravity, Jeans, or self-shielding criterion for SF). But with these parameters, $n^{\rm max} = 0.1\,{\rm cm^{-3}}$. Star formation is only possible in these simulations because the authors allow their SPH smoothing lengths (and equivalent hydrodynamic separations h_i) to be much smaller (by a factor 10) than $\epsilon_{\rm gas}$ (hence densities can reach $\sim 1000\,n^{\rm max} \approx n_{\rm crit}$) – but this means star formation (locally) has nothing to do with gravity! This is obviously unphysical, however, it is surprisingly common in the literature, see e.g. the Eris simulation in Guedes et al. 2011, who adopt $[\epsilon_{\rm gas},\,m_{i,\,1000},\,n^{\rm max}] = [180\,{\rm pc},\,20,\,0.1\,{\rm cm^{-3}}]$ but $n_{\rm crit} = 5\,{\rm cm^{-3}}$. We therefore consider two similar experiments: for **m10q**, we take $\epsilon_{\rm gas} = 20\,$pc ($n^{\rm max} = 10\,{\rm cm^{-3}}$) and for **m12i** take $\epsilon_{\rm gas} = 140\,$pc ($n^{\rm max} = 1\,{\rm cm^{-3}}$), but in both cases keep $n_{\rm crit} = 100\,{\rm cm^{-3}}$ and allow arbitrarily small h_i.

Remarkably, in both cases this appears to behave reasonably. This is because both galaxies have large turbulent motions that generate a broad distribution of density fluctuations – for gas which can cool to $\sim 10\,$K on reaching high densities, a turbulent $\delta v \sim 10\,{\rm km\,s^{-1}}$ should produce a lognormal distribution of densities with $\gtrsim 10\%$ of the mass exceeding $\sim 100\times$ the mean density (see e.g. Federrath et al. 2008). Indeed, smaller-scale simulations show most GMCs form by these fluctuations becoming so dense that they become self-gravitating, at which point they "detach" from the turbulent flow and collapse (see e.g. Hopkins 2013b). But with artificially large $\epsilon_{\rm gas}$ values, this "selection" via self-gravity cannot occur, so the gas will not automatically proceed to yet-higher densities: this means the properties of the cold, dense gas are incorrect, and sensitive (in a way they should not be) to $n_{\rm crit}$. But it does allow densities $> n_{\rm crit}$ to be reached. Since the system is ultimately feedback-regulated, as we will show in § 8, the galaxy comes to the same steady-state SFR where feedback balances inflow, by changing the turbulent and dense gas properties. For **m12i**, this self-regulation works well; for **m10q**, the too-large $\epsilon_{\rm gas}$ does significantly suppress the SFR and stellar mass (by a factor ~ 2.5) – and the SF is artificially bursty (more like one of our un-converged, low-resolution simulations). This is expected; in dwarfs, the turbulent Mach numbers are much lower (thermal gas support is more important), so the mechanism above requires some

additional, large-scale perturbation (it appears by-eye that gas accumulates until the disk goes globally gravitationally unstable, causing excessively large bursts in the galactic center instead of local fragmentation into clouds).

In both **m10q** and **m12i**, if we make $\epsilon_{\rm gas}$ still larger yet, even if we allow $h_i \rightarrow 0$, we eventually suppress any star formation. Of course, in this limit the entire galaxy is simply a warm gas-pressure supported disk or ball of gas, without effective self-gravity. In this limit, an orders-of-magnitude lower $n_{\rm crit}$, below the mean density of the galaxy, must be used, since sub-structure within the galaxy cannot be resolved. But then the SF treatment takes on an entirely different meaning (it is a sub-grid model for galaxy-scale fragmentation into GMCs, not a representation of how gas turns into stars within dense sub-clumps in GMCs). This is the limit in the large-volume simulations (EAGLE, Illustris) discussed in § 1.

4.2.6 *N-Body Heating Rates*

One might worry, instead, that N-body heating of gas/star particles by the dark matter (or stars, or other gas particles) could thicken the stellar disk, or inject spurious thermal energy/turbulence into the gas. Consider the worst-case hard-scattering scenario: two particles approach one another in a vacuum (no other particles nearby) and scatter near the particle centers, with encounter velocity $v_{\rm encounter}$. Since the duration of the encounter is short, the impulsive velocity change is $\Delta v = |\Delta \mathbf{v}| \sim G m_i/(h_i\, v_{\rm encounter})$ (a proper integration over the exact kernel shape used in the simulations gives a maximal deflection $\approx 1.3\, G m_i/(h_i\, v_{\rm encounter})$, which occurs for encounters with impact parameter $\approx h_i$). For gas, this is $\Delta v \sim 0.01\,{\rm km\,s^{-1}}\, m_{i,\,1000}^{2/3}\, (n/{\rm cm^{-3}})^{1/3}\, (v_{\rm encounter}/10\,{\rm km\,s^{-1}})^{-1}$. For dark matter, the particle masses are larger, but so are the softenings and typical encounter velocities since the dark matter has approximately isotropic dispersion at the virial velocity, $\sim 100\,{\rm km\,s^{-1}}$. So $\Delta v \sim 1.3\, G m_{\rm DM}/(h_{\rm DM}\, v_{\rm encounter}) \sim 0.017\,{\rm km\,s^{-1}}\, m_{i,\,1000}\, (h_{\rm DM}/20\,{\rm pc})^{-1}\, (v_{\rm encounter}/100\,{\rm km\,s^{-1}})^{-1}$, only slightly larger even when we adopt the smallest DM softening seen in our simulations.[18]

We can translate this to an "N-body heating rate," assuming each encounter adds to the velocity incoherently, with particle encounter rate $\sim v_{\rm encounter}/h_i$. Accounting for the number of baryons per gas particle ($\sim m_i/m_p$), this gives a heating rate (rate-of-change of kinetic energy) of:

$$\frac{Q_{\rm heat}^{gas-gas}}{{\rm erg\,cm^3\,s^{-1}}} \sim 2\times10^{-32}\, m_{i,\,1000}\, \left(\frac{10\,{\rm km\,s^{-1}}}{v_{\rm encounter}}\right) \quad (10)$$

$$\frac{Q_{\rm heat}^{gas-DM}}{{\rm erg\,cm^3\,s^{-1}}} \sim 6\times10^{-32}\, m_{i,\,1000}\, \left(\frac{100\,{\rm km\,s^{-1}}}{v_{\rm encounter}}\right) \left(\frac{\rho_{\rm DM}}{\rho_{\rm gas}}\right) \quad (11)$$

These should be compared to the physical heating/cooling rates of the gas, $\Lambda \sim 10^{-23}-10^{-22}\,{\rm erg\,cm^3\,s^{-1}}$ for typical ISM conditions; they are far smaller than almost any other source of error in the baryonic physics, even for the lowest-resolution simulations in our tests.[19]

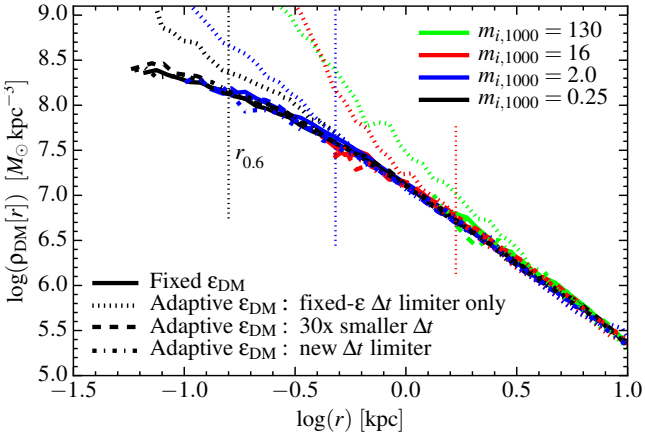

Figure 20. Dangers of using adaptive force softening for *collisionless* (non-gas) particles, without careful timestepping. We show the $z = 0$ mass profile in DM-only simulations (here our **m10v** run, but we repeated a subset of these tests in **m09**, **m10q**, **m11q**, **m11v**, and **m12i**), at a series of resolution levels (labeled). For our default (fixed-$\epsilon_{\rm DM}$) runs, we show the conservative $\sim$ 2000-particle Power et al. (2003) radius $r_{0.6}$, but plot the full profile down to our convergence radius $r_{0.06}$ – the profiles in all previous tests are well-converged outside this radius, independent of $\epsilon_{\rm DM}$, and we see the same in the fixed-$\epsilon_{\rm DM}$ runs plotted here. We compare our adaptive-softening implementation with three different timestep choices. If we use only the same timestep criteria employed for fixed-softening runs ("fixed-ϵ Δt limiter only"), the simulations clearly exhibit a spurious, non-converged "cusp" at the center indicative of integration errors that arise because particles "move through" one another in a timestep, preventing the adaptive terms from properly adapting (this cannot occur with gas, owing to the Courant condition). Lowering the timestep by a uniform factor of 30 resolves the problem, and produces excellent agreement with the fixed-ϵ runs, but at great computational expense. Our default implementation uses the "new Δt limiter" described in § 4.3.3 (Eq. 13) to control the timesteps with adaptive softening for collisionless particles: this also produces excellent agreement with fixed-ϵ runs.

4.3 Time Resolution

The time resolution in the simulations is set by the timestep Δt, which is always set according to the minimum of the various criteria in § 2.7. Together, these criteria ensure that gravitational and

[18] In fact, the Δv in the text above is actually a significant over-estimate when we use adaptive gravitational softenings, as particle-particle encounters are never "in a vacuum." While a particle a traverses a domain h within the kernel of particle b, it of course feels some acceleration. But provided there is actually mass being represented by b, this is completely physical! The numerical noise/error depends not on the absolute magnitude of the acceleration from b onto a, but on the *deviation* of the potential from b, owing to finite sampling, from the smooth potential that would be represented if we had infinite resolution. This is straightforward to estimate. Following Dehnen & Aly (2012), the acceleration from each particle is constructed assuming its mass distribution follows the kernel function; if we assume the correct background is a uniform density field, then discretize this into particles, it is straightforward to compute the fractional deviation from the correct (infinite-resolution) solution for different particle configurations within the kernel. For our standard cubic-spline kernel, this is essentially what is shown in Fig. 3 of Dehnen & Aly (2012); the typical deviation considering various different particle configurations is about $|\Delta\rho|/|\rho| \sim 0.005$. Thus the actual hard-scattering amplitudes are suppressed by a factor of ~ 100 from their already-small values.

[19] We have directly verified this in a series of numerical tests: setting up a periodic box with dark matter particles (matched to the particle mass, velocity dispersion, and space density of the cosmological simulations at the radii within the halo where the N-body heating in Eq. 11 is maximized), and an equilibrium, isothermal, non-radiative gas disk. In tests where the gas feels only DM, we measure $Q_{\rm heat}^{\rm gas-DM}$ a factor of a few *smaller* than Eq. 11; in tests where the gas feels self-gravity and is "stirred" (so some relative gas-gas motion exists), the heating contribution from $Q_{\rm heat}^{\rm gas-gas}$ is unmeasurably small.

fluid dynamics, as well as stellar evolution, are always explicitly time-resolved. Usually, the acceleration-based criterion $\Delta t < 0.2\,(h_i/|\mathbf{a}_i|)^{1/2}$ from Power et al. (2003) is the most demanding, for *both* gas and N-body particles (because the dense gas is highly super-sonic). Occasionally, however, in gas shock-heated by many SNe, the Courant criterion is most important.

In either case, the minimum timesteps reached in our simulations are $\approx 100\,$yr (see Table 3). This is reached regularly by some elements from redshifts $z \sim 0-2$, although always by a very small fraction of the total population at any given instant.

4.3.1 Standard Criteria & Results of Variations

We have tested variations in our standard time-stepping criterion, in a limited sub-set of simulations with both dark matter and baryons: the DM-only runs are run to $z = 0$, but baryonic runs are only integrated to $z = 4$, because inappropriate timesteps usually lead to catastrophic numerical instability which is evident quickly. Our conclusions are identical to canonical studies in the field. Like Power et al. (2003), we find that a coefficient smaller than $\alpha = 0.2$ in the timestep limiter $\Delta t < \alpha\,(h_i/|\mathbf{a}_i|)^{1/2}$ produces no appreciable gain in accuracy or convergence, but much larger values ($\alpha \gtrsim 0.5$) can seriously degrade mass profiles and orbit integration (e.g. angular momentum conservation); our gravity solvers are very similar, following Springel (2005), so this should not be surprising, especially for DM-only tests. Even though this acceleration-based criterion usually dominates for gas (recall $\mathbf{a}_i$ includes all accelerations, i.e. hydrodynamic and gravitational), a Courant-type condition is still necessary, with coefficients $\alpha > 0.8$ in $\Delta t < \alpha\,h_i/v_{\rm sig,\,i}^{\rm max}$ giving rise to numerical instability, while values $\alpha < 0.4$ (our preferred value) do not produce any obvious improvement. This is consistent with standard hydrodynamic tests and analytic numerical stability analysis (see Hopkins 2015 for discussion).

4.3.2 Stellar-Evolution Timestep Limits

Recall, we do not allow the timestep for star particles to exceed

$$\Delta t_* < \mathrm{MAX}\left(10^4\,\mathrm{yr},\ \frac{t_*}{300}\right) \tag{12}$$

where t_* is the age of the star, to ensure stellar evolution is time-resolved and the expected number of SNe (per star particle per timestep) is always < 1 at our production resolution. We have experimented with weakening this limiter, increasing it by a factor of ≈ 3 (which produced no measureable effect), or most radically using $\Delta t_* < 20\,\mathrm{Myr}\,m_{i,\,1000}^{-1}$ – for $m_{i,\,1000} \gtrsim 1$, this is equivalent to allowing as many as $\sim 10\,$SNe per timestep in an extremely young star particle. We did not see any significant difference running tests of **m10q** and **m12i** at low resolution (the difference in timestep is maximized at low resolution) to $z = 2$; however, we found that usually the number of SNe exploding at once was much smaller ($\sim 2-3$), because the youngest stars are in dense regions where other timestep limits (e.g. the acceleration criterion above) impose much stricter limits than $\sim$Myr. However, if we removed the limiter entirely ($\Delta t_* \rightarrow \infty$), we saw clear (albeit rare) pathological activity: for example lone star particles formed in poorly-resolved dwarfs in the outskirts of the high-resolution region or ejected via tidal interactions might be assigned extremely long dynamical timesteps and have a huge single-timestep injection of SNe, continuous stellar mass loss, and radiation (this would lead, for example, to a small number of gas elements with extraordinarily high metallicity, because everything was injected simultaneously from these pathological star particles).

4.3.3 Adaptive Force Softening for Collisionless Particles: The Need for Additional Timestep Criteria

As described in § 2.2, for a subset of our tests presented in § 4.2.4, we use adaptive softening for collisionless (DM and stellar) particles. Variations of these methods have been explored in a number of studies on collisionless, self-gravitating systems (Athanassoula et al. 2000; Price & Monaghan 2007; Bagla & Khandai 2009; Iannuzzi & Dolag 2011; Iannuzzi & Athanassoula 2013; Barnes 2012), and our specific implementation follows Price & Monaghan (2007), scaling the force softening to match the inter-particle spacing (DM with DM, stars with all baryons, gas with gas) $\epsilon_i = h_i$ and including the appropriate correction terms in the force equations to maintain exact momentum and energy conservation. We have validated our numerical implementation with all tests presented in Price & Monaghan (2007) and Iannuzzi & Dolag (2011).

However, introducing these adaptive softenings requires stronger timestep criteria for collisionless particles, especially in dense halo centers where DM particles on highly-radial orbits may "plunge" and interact with particles of widely-differing ϵ over the course of their orbits.

Fig. 20 shows the results of using adaptive softening in DM-only simulations (**m10v**) of varying resolution, *without* imposing any additional timestep limiter beyond what is described above for collisionless particles with *fixed* softening ϵ. Clearly, the central sub-kpc regions of the DM profile exhibit a spurious density "cusp" well in excess of the *converged* (much higher-resolution) results from constant-$\epsilon_{\rm DM}$ runs. We confirm that this feature is artificial by simply re-running the simulations enforcing a factor of 30 smaller timestep for the DM; this agrees well with the fixed-ϵ runs. We have also confirmed that both sets of runs maintain global conservation (as expected) – the error appears to be associated with local integration errors between particle neighbors when the timesteps are too large.

With fixed softening, there is no Courant-like condition required for collisionless particles: since the equations of motion depend only on the collective long-range forces (the gravitational potential), if the potential is sufficiently smooth ($|\mathbf{a}|$ is small), particles can safely "move through" one another (Monte Carlo-sampling the phase-space distribution function). But this is not true with adaptive softening (again, for collisionless particles), because the softening length (hence self-gravity) of a particle depends on the local neighbor configuration, introducing local correction terms that must be integrated smoothly as particles move through one another – otherwise these terms are under-sampled and effectively scatter the orbits.

For gas, this error does not occur because of the Courant condition. The solution for collisionless particles with adaptive softening is therefore straightforward: we should implement a similar timestep criterion. We require:

$$\Delta t_{\rm AGS}^{a} < 0.25\,\mathrm{MIN}\left\{\frac{1}{|\langle\tilde{\nabla}\cdot\mathbf{v}\rangle_a|}\ ,\ \frac{\epsilon_a}{v_{\rm sig,\,AGS}^{a}}\right\} \tag{13}$$

$$v_{\rm sig,\,AGS}^{a} \equiv \mathrm{MAX}_{(b:\,|\mathbf{x}|_{ba}<H_a,\,H_b)}\ \{|(\mathbf{v}_b-\mathbf{v}_a)\cdot\hat{\mathbf{x}}_{ba}|\} \tag{14}$$

$$\langle\tilde{\nabla}\cdot\mathbf{v}\rangle_a \equiv \frac{\sum_{b:\,|\mathbf{x}|_{ba}<H_a}(\mathbf{v}_b-\mathbf{v}_a)\cdot\nabla W(\mathbf{x}_{ba},H_a)}{\Omega_a\sum_b W(\mathbf{x}_{ba},H_a)} \tag{15}$$

Here $H_a \approx 2\,\epsilon_a$ (with $\epsilon_a = h_a$) is the domain of the nearest-neighbor search around particle a, and the sums over b represent sums over all interacting neighbors *of the relevant particle type* such that they contribute to defining the inter-particle spacing and softening length ϵ of particle a. By analogy to the Courant condition,

we define $v^{a}_{\rm sig,\,AGS}$ as the maximum approach/recession velocity of any neighbor within this interaction kernel – this requirement is simply that two particles cannot "cross" more than $\sim 1/2$ their relative softening lengths in a single timestep. The particle-divergence $\langle\tilde{\nabla}\cdot\mathbf{v}\rangle_a$ is defined exactly the same as the traditional SPH velocity divergence,[20] in such a manner that this requirement prevents ϵ_a from changing by more than $\sim 10\%$ within a single timestep. Because we use adaptive timesteps, we also enforce a "wakeup" condition identical to that used for the hydrodynamics (see Saitoh & Makino 2009; Durier & Dalla Vecchia 2012); specifically, if an "active" particle in a sub-step interacts with an "inactive" particle in a timestep >4 times larger ($v_{\rm sig,\,AGS}$ from its previous active step >4 times larger), the inactive particle is stopped from taking the larger timestep and moved to the smallest active timebin in the active hierarchy (see Hopkins 2015 for details). This prevents particles moving very rapidly from artificially moving "through" a particle with a long timestep.

Fig. 20 demonstrates that this timestep limiter cures the errors seen before. The added timestep criterion does add a significant cost to the DM-only run, though far less costly than the uniform factor ~ 30 smaller-timestep case, since only a small number of particles are affected at each time. In any case the agreement between adaptive softening with appropriate timesteps and fixed-softening runs is excellent down to $r_{0.06}$. We have also recently learned that other authors who have implemented adaptive DM softening following Price & Monaghan (2007) have reached the same conclusions and found it necessary to include similar timestep criteria to maintain numerical stability (V. Springel, private communication).

5 EFFECTS OF THE HYDRODYNAMIC METHOD

5.1 Finite-Volume Godunov Methods (MFM) vs. SPH

As noted in § 2.1, a range of idealized test problems demonstrate that our default, finite-volume MFM method for the hydrodynamics is significantly more accurate than SPH, at fixed mass resolution. However, given that gravity and feedback overwhelmingly dominate over pure hydrodynamic forces in the simulations here, does the hydrodynamic accuracy matter? Fortunately, GIZMO is an inherently multi-method code, so we can compare simulations with otherwise identical physics and numerics, replacing *only* the hydrodynamic solver.

Figs. 21, 22, 23, & 24 compare GIZMO simulations using both our (1) default (non-SPH) MFM method,[21] (2) various "flavors" of SPH.

[20] With adaptive softening for DM, in 3D, we *define* $\epsilon_a = h_a = (3N_{\rm eff}/4\pi)^{1/3}H_a$ such that, in 3D, $h_a^3\bar{n}_a = 1$ and $4\pi/3\,H_a^3\bar{n}_a = N_{\rm eff}$ (with $N_{\rm eff}=32$ for our standard cubic spline kernel), where $\bar{n}_a \equiv \sum W(\mathbf{x}_{ba}, H_a)$ is a kernel-averaged particle neighbor number density. With some straightforward algebra, this gives an *exact* discrete equation for the Lagrangian derivative of h_a, $Dh_a/Dt = -(h_a/3\Omega_a\bar{n}_a)\,\nabla\bar{n}_a\cdot\partial(\mathbf{x}_b-\mathbf{x}_a)/\partial t = -(h_a/3)\,\langle\tilde{\nabla}\cdot\mathbf{v}\rangle_a$, where $\Omega_a \equiv 1+(h_a/3\bar{n}_a)\,\partial\bar{n}_a/\partial h_a$ (for derivations, see Price 2012; Hopkins 2013a). So if we do not want the softening h_a to change by more than a factor $|(\Delta h_a)/h_a| < \alpha$ in one timestep, we require a timestep $\Delta t < 3\alpha/|\langle\tilde{\nabla}\cdot\mathbf{v}\rangle_a|$.

[21] We have also considered a limited comparison of the "meshless finite volume" (MFV) method in GIZMO. This differs from MFM only in a second-order mass-flux term, and gives nearly-identical results in test problems (Hopkins 2015; Hopkins & Raives 2016), so (unsurprisingly) the results are very similar to MFM. We prefer MFM as our "default" method because it maintains element masses, minimizing particle-splitting and reducing N-body noise.

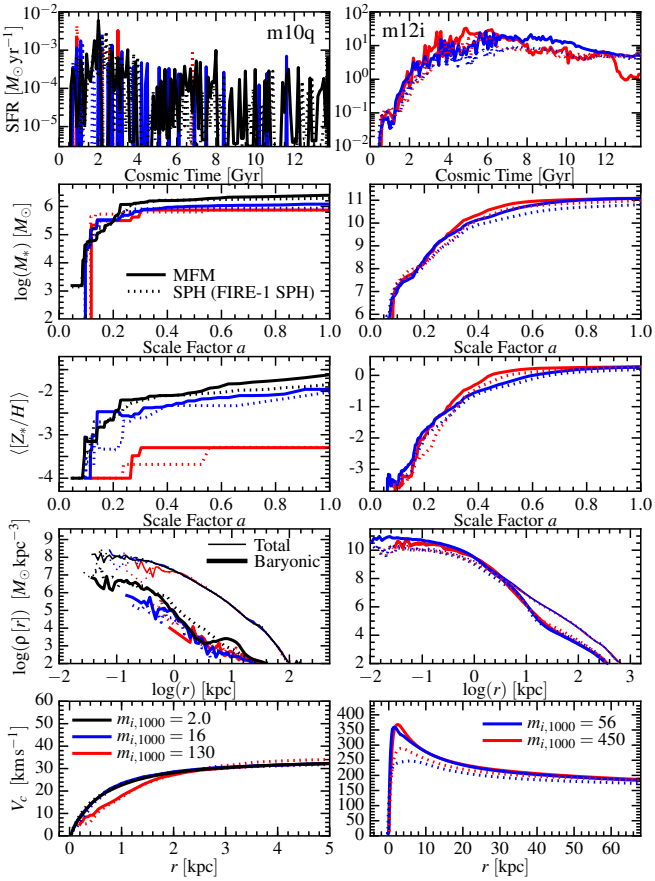

Figure 21. Effects of the method for solving the hydrodynamic equations, in cosmological simulations of both a dwarf (**m10q**; *left*) and MW-mass galaxy (**m12i**; *right*), as Fig. 4. We compare our default method in GIZMO – MFM (solid lines), a higher-order accurate, mesh-free finite-volume Godunov method – to the GIZMO implementation of smoothed-particle hydrodynamics (SPH) which was used for FIRE-1 (dotted lines), using more accurate switches for artificial viscosity and conduction and a larger smoothing kernel. GIZMO is a multi-method code so we can change the hydro solver while keeping *all* other physics and numerics identical here. We repeat each comparison at multiple resolution levels. For dwarfs, we see excellent agreement between MFM and SPH, at every resolution level. For MW-mass systems, on the other hand, FIRE-1 SPH predicts somewhat lower SFRs, stellar masses, and central V_c. The critical difference between dwarfs and massive galaxies appears to be the presence of the "hot gaseous halo" around massive systems (absent in dwarfs), which determines the cooling rate onto the galaxy and into which galactic winds propagate. Known issues in SPH can suppress fluid mixing in the halo, even with "state of the art" SPH formulations, which in turn leads to easier escape of winds, less efficient cooling, and (in turn) lower masses.

5.1.1 *"Flavors" of SPH*

There is an extensive literature of various SPH "flavors" which attempt to reconcile the differences between SPH and finite-volume methods like our MFM here (see e.g. Wadsley et al. 2008; Price 2008; Saitoh & Makino 2013; Hopkins 2013a; Read & Hayfield 2012; Hu et al. 2014; Rosswog 2014, and references therein). In this section we will consider three such flavors:

- **FIRE-1 SPH:** This is the SPH formulation used in FIRE-1, and described in Hopkins et al. (2014). Briefly, we use the "pressure-energy" formulation of SPH (P-SPH) from Hopkins (2013a) to eliminate the spurious "surface tension" error present in "density-energy" (D-SPH) formulations at contact discontinu-

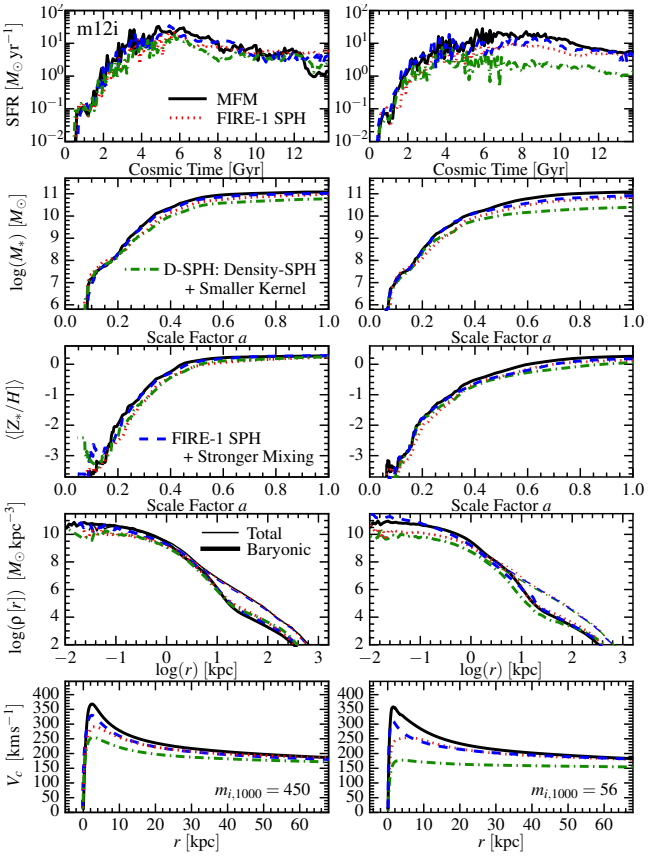

Figure 22. Effects of the hydrodynamic method in our cosmological **m12i** simulation at two resolution levels (*left* and *right*), as Fig. 21. Here we compare different "flavors" of SPH (see § 5.1.1): (a) the FIRE-1 implementation from Fig. 21, which uses the pressure-energy formulation and a larger smoothing kernel, designed to reduce fluid-mixing errors; (b) the FIRE-1 SPH model with "stronger mixing" – an explicitly increased thermal energy/entropy mixing term (larger "artificial conductivity"); (c) D-SPH: a simpler SPH implementation which uses the "density-energy" formulation of the equations and a smaller kernel (but still uses the higher-order artificial viscosity and conductivity switches of FIRE-1 SPH), which strongly suppresses the ability of the method to capture fluid mixing instabilities. The "FIRE-1 SPH + Stronger Mixing" run agrees extremely well with MFM; the "D-SPH" run much more strongly suppresses the SFR, stellar mass, and central V_c, in a manner which does appears to *diverge* with resolution (consistent with the fact that the SPH errors are zeroth-order). This clearly demonstrates that the effects of SPH on fluid mixing physics dominate the differences between runs in Fig. 21.

ities; a higher-order kernel (the quintic spline; with ~ 64 effective neighbors instead of ~ 32 for the cubic spline in MFM) to reduce zeroth-order SPH errors (see Dehnen & Aly 2012; Zhu et al. 2014); more accurate gradient estimators using the moving-least-squares approach (as in García-Senz et al. 2012; Rosswog 2014); higher-order switches to minimize the artificial viscosity following Cullen & Dehnen (2010) with further improvements described in Hu et al. (2014) and Hopkins (2015); and added "artificial conductivity" terms to allow entropy/thermal energy diffusion similar to artificial viscosity (Wadsley et al. 2008; Price 2008). These improvements to SPH are designed to improve behavior in fluid-mixing and shock capturing, while reducing noise and artificial numerical diffusivity away from shocks (see references above).

Readers interested in more details of the SPH method here

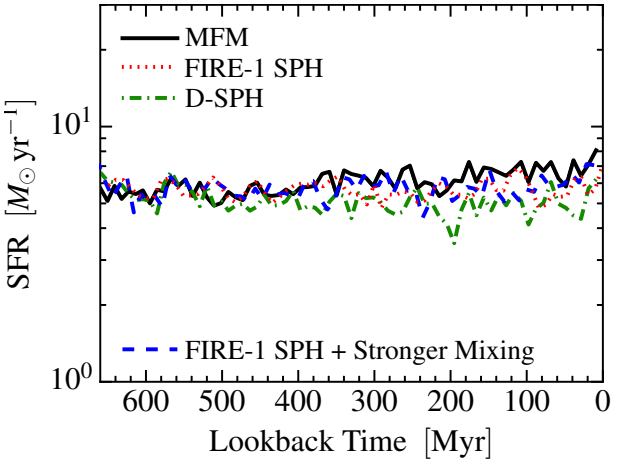

Figure 23. Effects of the hydrodynamic method on the SFR at low-redshift of a re-started MW-mass (**m12i**; $m_{i,1000} = 56$) simulation as Fig. 10. We compare MFM and various SPH "flavors" as Fig. 21. For fixed initial conditions (early times) the SFR is feedback-regulated and identical. The two slowly diverge by $\sim 30\%$ over $\sim$ Gyr as the SPH flavors with less accurate fluid-mixing treatments are able to more easily eject gas from the galaxy for spurious numerical reasons. However the effect is small compared to the cosmologically time-integrated effects seen in Fig. 21. This is consistent with the idea that the mixing and re-cycling of these outflows, and subsequent hot gas cooling (which occurs on timescales of order the Hubble time), not the *generation* of outflows or self-gravitating fragmentation and star formation in the dense galactic disk, is the dominant reason for the difference between certain SPH flavors and MFM in massive galaxies in Fig. 21.

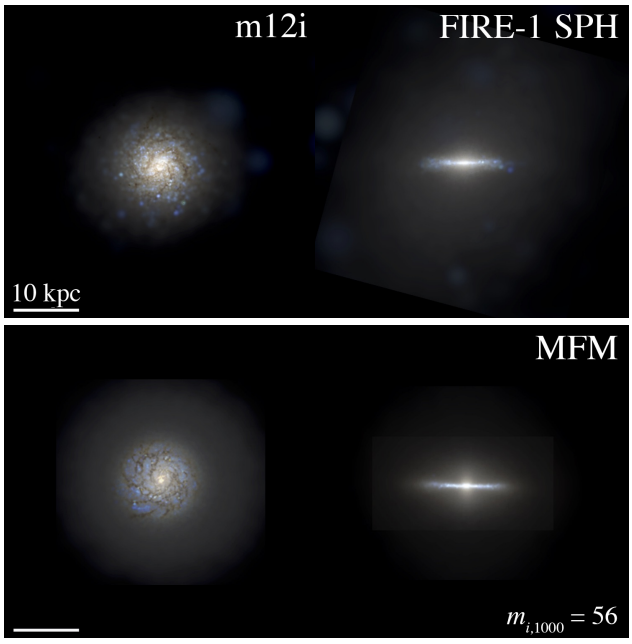

Figure 24. Mock images, as in Fig. 3, showing the effects of the hydrodynamic solver on the $z = 0$ morphology of the simulated galaxies, for the highest-resolution ($m_{i,1000} = 56$) MW-mass (**m12i**) systems in Fig. 21. Dwarfs are not shown, as they have irregular morphologies independent of the hydrodynamics or resolution (as expected). The large-scale visual morphology is similar, although the MFM simulation exhibits a slightly more compact, thinner disk and more well-ordered (less noisy) spiral structure. The major differences in stellar mass and central circular velocity are not obvious in the stellar morphology. The same is true for all SPH flavors in Fig. 22.

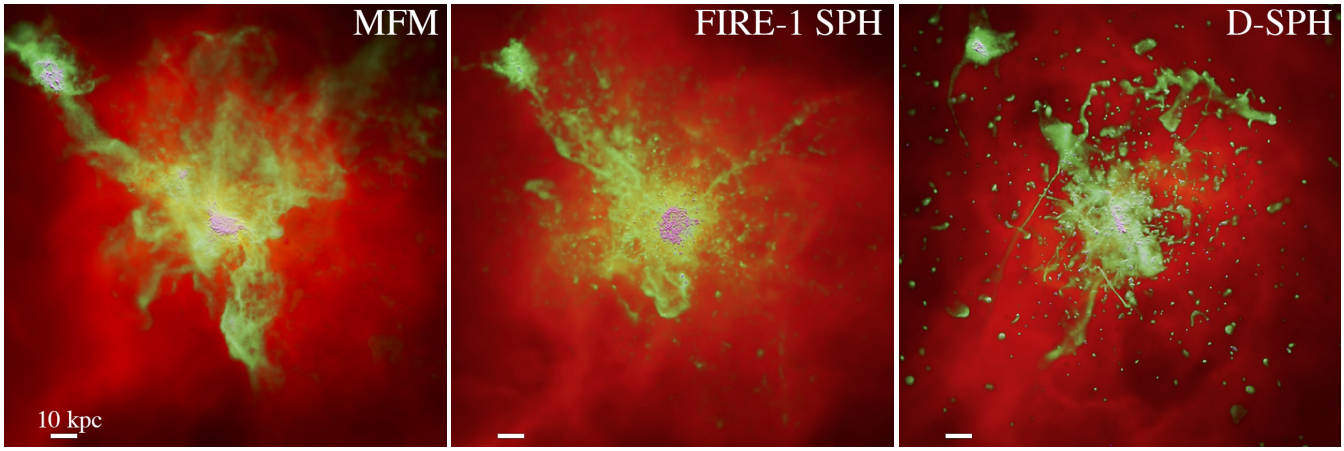

Figure 25. Gas morphology around the galaxy in the **m12i** simulations (with $m_{i,1000} = 56$) from Fig. 21, at $z = 0.9$ in 200 kpc boxes, with different hydrodynamic methods (as Fig. 21; labeled). Images are logarithmically-weighted surface-density projections, with red/green/magenta showing hot ($> 10^6$ K), warm ionized ($\sim 10^4 - 10^5$ K), and cool neutral (< 8000 K) gas. The most dramatic differences appear in the CGM; we choose a time where a merger has triggered violent outflows of cool gas to maximize these differences. In some less-accurate SPH formulations, such as "D-SPH" here (*right*; density-SPH with smaller smoothing kernel), the outflow has broken into the well-known artificial "SPH blobs" that result from errors treating fluid mixing interfaces. In our non-SPH, finite-volume MFM method (*left*), these are absent. The FIRE-1 implementation of SPH ("P-SPH" with larger kernel) was specifically formulated to reduce the fluid-mixing errors in SPH; it dramatically reduces but does not completely eliminate the "blobs" (FIRE-1 SPH+Stronger Mixing closely resembles this, but with slightly fewer "blobs"). Also, in MFM the hot halo gas is more compact/dense, with a sharper shock front, while in SPH it is lower-density with an extended boundary, owing to difficulties in shock-capturing and numerical dissipation/mixing. These effects in SPH make it "easier" for cold winds to be ejected and lower the CGM gas cooling rate, explaining the suppression in SF in Figs. 21-23.

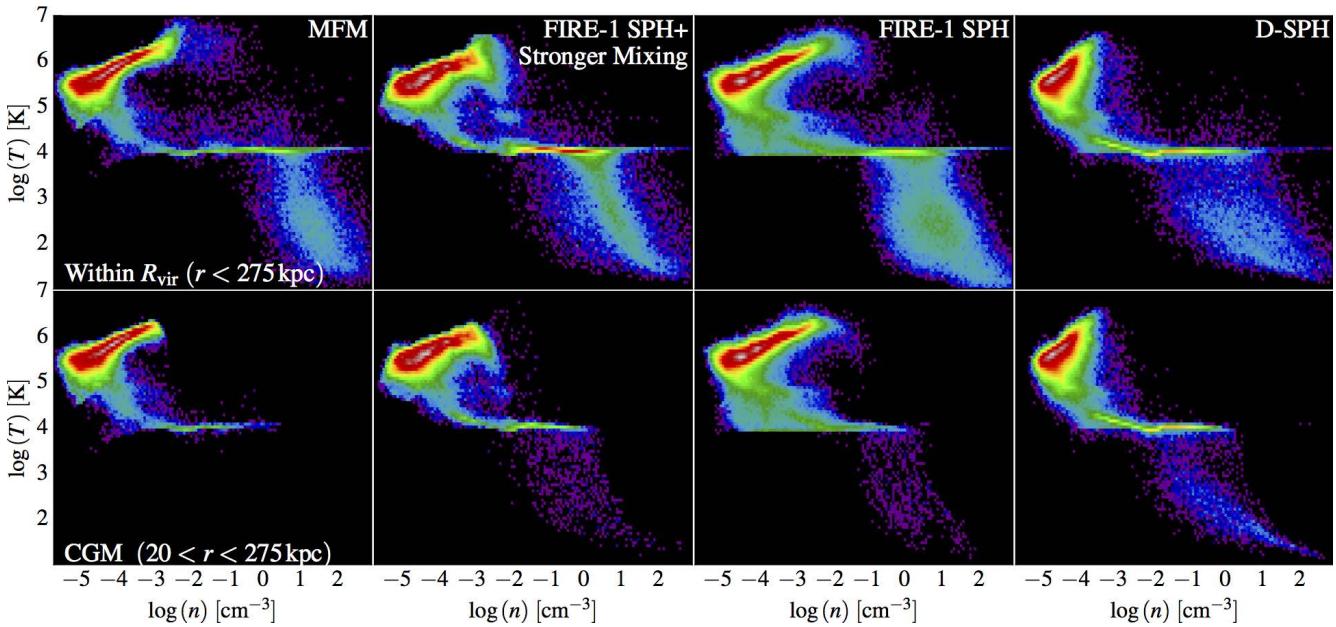

Figure 26. Comparison of hydrodynamic methods (MFM and SPH "flavors" from Fig. 22) in the temperature-density phase diagram (at $z = 0$) in the MW-mass (**m12i**, with $m_{i,1000} = 56$) systems in Figs. 21-25. We compare all gas inside the virial radius (*top*) and gas within the CGM (excluding the galaxy; *bottom*). Colors are a mass-weighted heat map with the density of gas mass in the space shown increasing logarithmically from black-blue-green-yellow-red-white. The "D-SPH" (simpler, less accurate fluid-mixing) flavor produces no hot halo gas with densities $n \gg 10^{-4}$ cm^{-3}, owing to a combination of poor shock-capturing, numerical dissipation, and artificial "ease" with which cold "blobs" in shredded galactic winds escape the halo (instead of shocking); there is also a substantial population of gas with temperatures $T \sim 10 - 1000$ K and densities $n \sim 10^{-2} - 10^3$ cm^{-3} in the CGM, from the same "blobs." The more accurate fluid mixing treatment in the FIRE-1 SPH mostly eliminates the cold CGM blobs and restores hot dense gas – but there is still a significant amount of gas at $\sim 10^4$ K even at very low densities which is likely spurious and dissappears when we add "stronger mixing," in good agreement with our MFM simulations.

should consult the public version of GIZMO: this is the default SPH implementation.

- **FIRE-1 SPH + Stronger Mixing:** A common feature in most of the SPH flavors above is the use of an "artificial conductivity" term (as in FIRE-1 SPH) which allows for diffusion of entropy between particles: without this, entropy becomes "particle-locked" leading to artificial resolution-scale discontinuities that suppress fluid mixing instabilities.

With these "artificial diffusion" terms in SPH, there is considerable freedom to adjust the form and normalization of the diffusivity. While the prescriptions are usually tuned to give some desired balance of accuracy on a mix of different idealized test problems, there

is usually no single "correct" prescription. We therefore consider the effects of adjusting this diffusivity term to give "stronger mixing" in SPH, in order to eliminate the spurious "blobs" discussed below. Specifically, we take the artificial conductivity in the GIZMO SPH implementation to be:

$$\frac{d\,E_a^{\rm SPH}}{dt} = \sum_b \alpha_{ab}\,\tilde{v}_{ab}^{\rm sig}\,\frac{m_a\,m_b}{\bar{\rho}_a+\bar{\rho}_b}\,(u_a-u_b)\,\tilde{W}_{ba} \tag{16}$$

$$\tilde{W}_{ba} \equiv \frac{\partial W(|\mathbf{x}_{ba}|,h_a)}{\partial|\mathbf{x}_{ba}|} + \frac{\partial W(|\mathbf{x}_{ba}|,h_b)}{\partial|\mathbf{x}_{ba}|} \tag{17}$$

$$\tilde{v}_{ab}^{\rm sig} \equiv c_{s,a}+c_{s,b}-3\,(\mathbf{v}_b-\mathbf{v}_a)\cdot\hat{\mathbf{x}}_{ba} \tag{18}$$

$$\alpha_{ab} \equiv \Theta(\tilde{v}_{ab}^{\rm sig})\,\frac{|P_a-P_b|}{P_a+P_b}\begin{cases}\dfrac{\alpha_a^V+\alpha_b^V}{8} & (\text{"Default SPH"'})\\ 1 & (\text{"Stronger Mixing"'})\end{cases} \tag{19}$$

$$\Theta(\tilde{v}_{ab}^{\rm sig}) \equiv \begin{cases}0 & (\tilde{v}_{ab}^{\rm sig}\le 0)\\ 1 & (\tilde{v}_{ab}^{\rm sig}>0)\end{cases} \tag{20}$$

where $\mathbf{x}_{ba}\equiv\mathbf{x}_b-\mathbf{x}_a$ is the particle separation, m mass, $\bar{\rho}$ the SPH-estimated density, W the SPH smoothing kernel as a function of $\mathbf{x}_{ba}$ and smoothing length h_a, c_s the sound speed, $\mathbf{v}$ the velocity, P the pressure, and α_a^V is the artificial viscosity coefficient defined in Hopkins 2015, Appendix F2, Eq. F16-F17. This α^V varies between $0.05-2$ depending on the velocity divergence $\nabla\cdot\mathbf{v}$ and its time-derivative, reaching large values when particles approach increasingly rapidly ($\nabla\cdot\mathbf{v}<0$ and $d[\nabla\cdot\mathbf{v}]/dt<0$) and decaying rapidly otherwise.

The difference between our "default" FIRE-1 SPH and "Stronger Mixing" simulations, therefore, is that "FIRE-1 SPH" (by design) only allows for mixing/diffusion of entropy and thermal energy when particles are effectively inter-penetrating or "move through one another." In the "Stronger Mixing" case, we make the effective diffusivity/mixing stronger in those cases (by a factor ≈ 4), but more importantly, we allow mixing in shear flows. The latter case has been shown to improve the accuracy of SPH in modeling fluid mixing instabilities such as the Kelvin-Helmholtz instability (Price 2008) and more accurately represents the "unresolved turbulent mixing" terms advocated in SPH by Wadsley et al. (2008); it is also numerically closer to the implementation advocated by Read & Hayfield (2012), who have argued this gives better convergence modeling cores of cooling-flow halos (Power et al. 2013). On the other hand, the "Stronger Mixing" implementation tends to "smear out" shock fronts and Keplerian shear flows much more significantly, further reducing the "effective resolution" of SPH (Hu et al. 2014; Few et al. 2016) – this is why it was not our "default" approach in FIRE-1. Note, though, that even in the "Stronger Mixing" case, there is still a "switch" α_{ab} which suppresses diffusion in supersonically receding flows and sharp phase discontinuities.

- **Density-SPH + Smaller Kernel (D-SPH):** Alternatively, we can compare an implementation of SPH where we *remove* some of the improvements used in our FIRE-1 SPH method. Specifically, in these simulations, we use the "density-energy" SPH (D-SPH) formulation instead of the "pressure-energy" SPH formulation, which introduces a non-convergent (sub-zeroth-order) error that has the functional appearance of a surface tension force at phase discontinuities (see Saitoh & Makino 2013; Hopkins 2013a). We also reduce the size of the SPH smoothing kernel, using a cubic spline with ~ 32 effective neighbors; this increases the zeroth-order SPH errors and reduces the ability of the code to capture fluid-mixing instabilities. This is the "T-SPH" formulation studied in Hopkins (2015). We retain all other aspects of FIRE-1 SPH (specifically the higher-order gradient, artificial viscosity, and artificial conductivity estimators). We consider this implementation because it represents some flavors of SPH still used frequently in the literature, and because it allows us to specifically highlight the effects of SPH errors that suppress fluid mixing at contact discontinuities.

5.2 (Weak) Effects in Dwarfs & the ISM of Massive Galaxies

Fig. 21 shows that at all resolution levels, our dwarf galaxy simulations are nearly identical in SPH and MFM methods (at all redshifts). It also clearly shows that SPH and MFM agree well on the properties of the main progenitor of massive (MW-mass) galaxies *while the progenitor is still a dwarf* with stellar mass $M_* \lesssim 10^{10}\,M_\odot$ at redshifts $z\gtrsim 1.5-2$. This holds for all properties plotted and all others we have examined: the distribution of mass in different phases, CGM gas morphology, covering fraction of absorbers, metallicity gradients, galaxy rotation, and more.

This is not surprising: such galaxies are in the "cold accretion" regime where there is little or no "hot halo" of virial shock-heated gas surrounding the galaxy (e.g. Kereš et al. 2005). Rather, gas falls onto galaxies on a free-fall time before being expelled by feedback. As such the properties are simply regulated by a combination of feedback and self-gravity; subtleties of fluid mixing instabilities (where SPH differs from other hydrodynamic methods) are unimportant. This is consistent with previous studies in FIRE-1 using different "flavors" of SPH, which had no effect on dwarf galaxy properties (Hopkins et al. 2014).

Figs. 21-23 show that for MW-mass galaxies, which do have hot halos, the qualitative behavior is similar in MFM and SPH, but some significant quantitative differences appear (see below). However, Fig. 23 shows that if we re-start an identical initial condition in MFM and SPH at low redshift, the SFRs are initially identical, and only slowly drift apart (here by just $\sim 20\%$ over $\sim 1\,$Gyr). Moreover Fig. 24 shows that the visual galaxy morphologies in SPH and MFM are very similar.

These results are consistent with a series of studies of *non-cosmological* simulations of MW-mass galaxies, comparing different SPH flavors (Hopkins et al. 2012b, 2013c) and SPH versus moving-mesh codes (Hayward et al. 2014). These studies showed that *within galaxies* known SPH errors have little effect on predictions, since the turbulence is primarily super-sonic, cooling is fast compared to dynamical times, and phase structure is primarily driven by gravitational collapse and feedback – all limits where SPH performs well (see Kitsionas et al. 2009; Price & Federrath 2010; Hopkins 2015).

5.3 (Significant) Effects in the CGM of Massive Galaxies' "Hot Halos"

However, despite the similarity in the SFRs in re-started massive galaxies and non-cosmological simulations, Figs. 21-22 clearly show that the *cosmological* SFHs of MW-mass systems diverge between some SPH flavors and MFM, as they reach stellar masses $M_*\gtrsim 10^{10}\,M_\odot$, corresponding to halo masses $\gtrsim 10^{11.5}\,M_\odot$. In the D-SPH runs (which maximize the difference), the final stellar mass is suppressed relative to the MFM run by a factor of ~ 1.5 (low resolution) or ~ 3 (intermediate resolution), giving a correspondingly smaller peak circular velocity.

Similarly, Davé et al. (2016) compare low-resolution cosmological large-volume simulations using a simplified, non-FIRE subgrid ISM and feedback model and MFM and "P-SPH" (FIRE-1 SPH) methods in GIZMO. Over the mass range simulated ($M_*\sim 10^9-10^{12}\,M_\odot$), the stellar mass functions agree very well at $z>1$, but then begin to differ, with SPH tending towards smaller masses

by ~ 0.2 dex by $z = 0$. Yet another study using still different ISM and feedback models in GIZMO presented in Zhu & Li (2016), who see a similar "divergence" between MFM and some SPH flavors.

Not coincidentally, this epoch where SPH and MFM diverge corresponds precisely to the formation of the "hot halo" (where virial-shocked gas has a cooling time longer than the dynamical time and establishes a steady-state atmosphere within the halo; see Kereš et al. 2005). In fact, many studies have shown that much, if not most, of the fuel supply for massive galaxies at later times owes to recycled wind material trapped in the CGM in these hot halos (see Oppenheimer et al. 2010; Davé et al. 2011b; Faucher-Giguère et al. 2011; Christensen et al. 2016; Muratov et al. 2016; Anglés-Alcázar et al. 2016). But this requires following the interaction of multi-phase winds moving sub-sonically or trans-sonically through a pressure-supported medium – precisely the regime where the known SPH errors are most problematic (O'Shea et al. 2005; Agertz et al. 2007; Read et al. 2010; Bauer & Springel 2012; Sijacki et al. 2012; Torrey et al. 2012).

Fig. 25 shows a visual comparison of the gas morphology and phase structure in the CGM at redshifts $z \sim 1$, where the differences between methods are most apparent. In D-SPH – which, we emphasize, is a formulation of SPH known to introduce larger errors compared to our "FIRE-1 SPH" – cool/cold gas in the CGM is primarily locked into numerically spurious "blobs." These are a well-known result of numerical errors in SPH, specifically its inability to capture multi-phase fluid mixing interfaces near the resolution scale (see Agertz et al. 2007; Kereš et al. 2012; Saitoh & Makino 2013; Power et al. 2013; Hu et al. 2014; Few et al. 2016). This causes inflows and outflows to "shred" into blobs, rather than properly mixing. The FIRE-1 "P-SPH" formulation is specifically designed to remove the specific "surface tension" error in SPH, which allows the blobs to be self-insulating and very long-lived. This reduces, but does not completely eliminate, the spurious "blobs." We note that all SPH variants shown include the "artificial conductivity" (entropy diffusion term) above, which has been claimed to eliminate such spurious structures (see Price 2008; Wadsley et al. 2008; Read & Hayfield 2012); we will discuss this further below.

Fig. 25 also shows that in SPH, the halo is more extended with more diffuse boundaries. This is also related to well-known SPH errors, specifically grid-scale heating/noise and poor shock-capturing, as well as excessive numerical dissipation of sub-sonic and trans-sonic turbulence (see e.g. Bauer & Springel 2012; Sijacki et al. 2012; Kereš et al. 2012). Both Springel (2010) and Hopkins (2015) show that SPH (even with state-of-the-art artificial viscosity switches from Cullen & Dehnen 2010) produces larger velocity noise around cosmological shocks (compared to MFM or moving-mesh methods), damps turbulence excessively below Mach numbers $\mathcal{M} \sim 1$, and requires $\sim 8-10$ inter-particle spacings ($\sim 30-50$ kpc at the virial radius for the low resolution in Fig. 25) to fully capture shock jumps.

The differences are even more striking when we examine the $z = 0$ temperature-density phase diagram in Fig. 26. While both SPH and MFM produce a significant amount of warm ($\sim 10^5$ K) and cool ($\sim 10^4$ K) CGM gas, we see clearly that in D-SPH there is a large amount of cold ($T \ll 10^4$ K) gas which has survived in the CGM being "protected" in cold, dense lumps, despite having extremely short physical mixing times (without magnetic fields – not included here – to "protect" the clouds). Also, in D-SPH, there is almost no hot, intermediate-density gas ($T \gtrsim 10^6$ K with $n_H \gtrsim 10^{-4}\,\mathrm{cm}^{-3}$), owing to the failures of shock-capturing (both virial and SNe shocks either being missed, as winds remain clumpy, or smeared-out).

Together, these errors suppress cooling in some SPH flavors, especially from hot gas in more massive halos (see Kereš et al. 2012; Torrey et al. 2012; Zhu & Li 2016), and they make it relatively "easier" for outflows to be ejected from the galaxy (because they are not properly shocking and/or mixing with surrounding halo gas). The artifacts mostly vanish in our more accurate FIRE-1 SPH treatment, although there is still some very low-density, cool gas which appears to be associated with "blob-like" structures. On the other hand, at low resolution, MFM features no spurious "blobs," and a denser hot halo with sharper shock boundaries – in much better agreement with the results of high-resolution moving-mesh or AMR codes, and completely expected from behavior in idealized test problems. As with all Godunov methods, there is almost certainly some numerical *over*-mixing at low resolution in MFM; this is part of the reason for the reduced stellar mass at higher resolution (Fig. 8), although that also owes to better resolution of SNe remnants.

Perhaps most troubling, however, in the simplest SPH formulations, for example the D-SPH flavor here, the "blobs" and associated under-mixing errors do not converge away at higher resolution, but persist at the resolution scale and actually reach larger density contrasts and contain *more* mass at higher resolution (this is because they owe to sub-zeroth-order errors; see Agertz et al. 2007; Read et al. 2010). We confirm this in all our tests here – this is why our D-SPH runs at MW-mass deviate *more* dramatically from MFM at high-resolution, and do not appear to be converging. We stress, however, that published FIRE-1 simulations used an improved SPH method that significantly reduced blob artifacts.

5.4 Comparing SPH "Flavors": Adding Diffusion Resolves Differences in the CGM

Fig. 22 also shows the results of using the "FIRE-1 + Stronger Mixing" SPH flavor instead of our default "FIRE-1 SPH" in the cosmological **m12i** simulation. With this enhanced mixing, the results agree extremely well with our MFM simulations. As expected, this "Stronger Mixing" SPH agrees even more closely with our MFM runs at dwarf masses. In contrast, as noted above, the D-SPH simulations using the "density-energy" formulation and a smaller kernel (both of which increase the SPH errors on fluid-mixing problems) show the largest disagreement with MFM.

This clearly demonstrates that the MFM-SPH differences, where present, owe to the treatment of fluid mixing, particularly how winds in "hot halos" do or do not mix and recycle onto the galaxy. This is a challenging numerical problem, and is almost certainly incompletely resolved in *any* numerical galaxy formation simulation. Therefore it is difficult to say which method is "more correct" at low resolution – rather we simply urge caution in interpreting these results at present. The major caveat which must be borne in mind for SPH is that, if the answer depends on an arbitrarily adjustable parameter (e.g. the artificial conductivity), and is non-convergent (as we find here), then SPH has unfortunately limited predictive power – we are forced to calibrate the SPH method at each resolution level to calculations with other codes, rather than simply increase the SPH resolution directly and trust that the errors should converge away (for more discussion, see Zhu et al. 2014; Zhu & Li 2016).

This also indicates that the large central V_c "spikes" we see in some of our low-resolution MFM runs, which are also sensitive to resolution, are fundamentally related to the same wind-mixing physics in hot halos.

For additional systematic comparison of "improved" SPH and MFM methods in recent cosmological hydrodynamic simulations

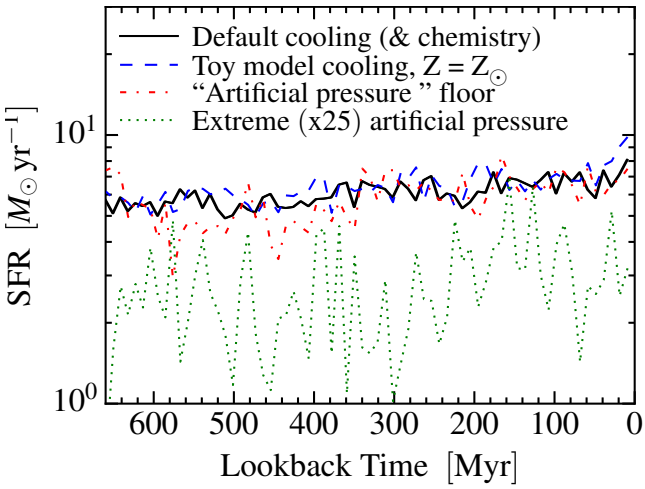

Figure 27. Effects of cooling physics details (§ 7) and artificial "pressure floors" (§ 6) on a restart of our MW-mass **m12i** simulation near $z = 0$, as Fig. 10. Our default model is full-physics, with no "artificial pressure." **(1)** *Toy Model Chemistry/Cooling:* We replace our default, detailed cooling physics with the toy model in § 7 which puts all gas on a single, solar-metallicity cooling curve, artificially turns off self-shielding (preventing cooling to $T \ll 10^4$ K), and removes the requirement that SF be in molecular gas. Because the cooling time is much faster than the dynamical time in all cases, details of the cooling functions have almost no effect on SF. **(2)** *Artificial Pressure Floor:* We add an artificial numerical "pressure floor" as described in § 6, with "modest" values designed to artificially inflate the thermal Jeans length to always be equal to four softening lengths. For these values this produces some spurious artifacts in e.g. the GMC and star cluster mass function (and violates energy conservation), but does not change any galaxy-scale results, since super-sonic turbulence and resolved collapse from much larger scales dominate SF. **(3)** *Extreme Artificial Pressure:* We now inflate the artificial pressure so the thermal Jeans length is always > 20 softening lengths. This is far larger than reasonable and makes the entire disk thermally-supported with $Q \gg 1$, shutting down substructure. Star formation only occurs when the disk becomes globally unstable leading to "mini-bursts" – the dynamics are clearly unphysical.

of galaxy formation, we refer readers to Zhu & Li (2016). Although the simulations there use a completely different treatment of feedback, cooling, and star formation, many of the conclusions – most importantly regarding the effects of fluid mixing in different methods – are identical.

6 PROBLEMS OF "ARTIFICIAL PRESSURE" TERMS

In order to follow ISM structure and star formation self-consistently, we specifically avoid the use of artificial or effective equations of state in our simulations. We emphasize that in FIRE-2 we do *not* adopt, as some previous studies (including FIRE-1) have, an additional "artificial pressure," numerically forcing the Truelove et al. (1997) criterion to have some value. In such approaches, an additional pressure of the form $P_{\rm artificial} \sim G\,(N\,h\,\rho)^2/\gamma$ is added to the equations of motion, where h is the resolution (local *gravitational* softening), ρ the density, and N is approximately the ratio of the Jeans length to h (after the artificial pressure is added) – this artificially moves the thermal Jeans mass to an unphysical, but easily-resolved, scale. In Lagrangian codes, this forms an adiabat with $P_{\rm artificial} \propto \rho^{4/3}$.

A common misconception is that this "prevents artificial fragmentation," but this is not true (see Teyssier 2015); in cases where $P_{\rm artificial}$ dominates, the medium is *by definition* Jeans unstable, so fragmentation (usually down to unresolved scales) is physically correct. Rather, what the artificial pressure does is move the fragmentation scale up to a larger scale, set by $\sim N\,h$, and suppress any fragmentation on smaller scales. This is numerically convenient in some circumstances, when one does not wish to follow dense fragments, for example. However, given our star formation prescription, it is clearly not appropriate to suppress fragmentation. Indeed, a cold (weakly pressure-supported), Jeans-unstable, self-gravitating region, with a Jeans scale that is small (stellar-mass scale), is *precisely* the sort of region which should (physically) be identified as star forming, and which our criteria attempt to correctly identify. The un-resolved fragmentation to small scales is already well-handled, in our method, by assigning the mass to a "star particle" which represents an aggregate of stars (formed via the un-resolved fragmentation to solar-mass scales).[22]

Furthermore, "artificial pressure" prescriptions are designed for entirely thermally-supported disks; it is not clear how to properly generalize them when turbulent, radiation, magnetic, or cosmic ray pressure dominates, but it is at least clear that applying the same prescription is likely to grossly over-estimate the "desired" pressure support (see e.g. Myers et al. 2013; Lee et al. 2014). Moreover, Kratter & Lodato (2016) note that in Lagrangian methods such as ours, capturing fragmentation properly and avoiding numerical errors from un-resolved fragmentation does *not* require additional pressure, but simply resolving the (GMC-scale) Toomre mass. In a rigorous study using Lagrangian methods in self-gravitating disks (the case of interest here), Nelson (2006) show that so long as (1) the Toomre mass is resolved (our own criterion in § 4.1.3), (2) the gas disk scale-height is resolved (identical to the Toomre criterion if $Q \sim 1$), and (3) fully-adaptive gravitational softenings are used for gas ($\epsilon_{\rm gas} = h_i$), the fragmentation cascade converges accurately and is numerically stable (resolution will truncate the lower-limit of the cascade, but all the power is on large scales, just as in turbulent motions) – no "artificial pressure" is necessary for good numerical behavior. This appears to be confirmed by the excellent convergence we find in the GMC mass functions and SFRs (Figs. 10 & 13). Chiaki & Yoshida (2015) implement particle-splitting to always satisfy the Truelove et al. (1997) condition, instead of artificial pressure, and similarly conclude no pressure (or splitting) is actually necessary for stability or convergence in Lagrangian methods.

Most problematically, the introduction of these "artificial pressure" terms violates energy conservation. Even if they only appear in the momentum equation (rather than being inserted as internal

[22] In fact, it is worth noting that Truelove et al. (1997) (or similarly Bate & Burkert 1997) never advocated using an "artificial pressure." Rather, their conclusion was that in order to *explicitly* resolve fragmentation *down to the thermal Jeans mass scale* (the brown dwarf scale, $\sim 0.1\,M_\odot$, in the cold ISM), one needed to maintain $\gtrsim 4$ resolution elements per Jeans mass. This can be accomplished by using the Truelove et al. (1997) criterion as a local *refinement* criterion (in an AMR sense), but that necessarily *requires* actually following fragmentation to sub-solar mass scales (standard practice in AMR simulations of individual proto-star formation). Alternatively, Federrath et al. (2010) and many subsequent authors (e.g. Padoan & Nordlund 2011; Myers et al. 2013; Gong & Ostriker 2013; Federrath 2015; Skinner & Ostriker 2015) pointed out that one can accurately use a Truelove et al. (1997)-type criterion to determine when local fragmentation is unresolved, and assign the mass to sink particles instead of following it explicitly (now standard practice in GMC and ISM-scale simulations) – this is, of course, our star formation prescription. Adding artificial pressure, on the other hand, will formally ensure the Jeans resolution criterion is met, but also guarantees the *wrong* Jeans mass, and therefore an incorrect solution.

energy, which would "restore" energy lost to cooling), they provide a potentially infinite source of "*PdV* work."

We therefore strongly discourage the use of these "artificial pressure" terms in any star-formation simulation which seeks to use sink-particle-type criteria (self-gravitating, cold, Jeans-unstable) for identification of star formation. An artificial pressure would be appropriate if, instead, one wished to assign some empirical large-scale star formation law to the gas (e.g. "putting in the Kennicutt-Schmidt relation by hand").

That said, we have checked whether adding such a prescription in the simulations makes a qualitative difference to our results; in Fig. 27, we implement the standard prescription above with $N = 4$. For $N \sim 1-5$, the difference in bulk galaxy properties (SFR, sizes, stellar masses) are small. This is because, for $N = 4$, the artificial pressure will stabilize an otherwise unstable spherical overdensity with radius $R \lesssim 30\,{\rm pc}\,m_{i,1000}^{1/3}\,(n/10\,{\rm cm^{-3}})^{-1/3}$ (where $m_{i,1000}$ is the particle mass in units of $1000\,M_\odot$); so for the resolution and typical gas densities inside the central ~ 4 kpc (the half-light radius) in Fig. 27, the largest GMCs ($R \gtrsim 100$ pc), which dominate the galaxy-integrated SFR, can still collapse. Manuel et al. (2016) reach similar conclusions (the added terms have little effect) in GMC-scale simulations. However, as expected there are several clearly un-physical artifacts introduced by an artificial pressure term, including: spurious noise in the low-temperature gas thermal properties (since it feels much stronger artificial numerical pressure), "bumps" in the mass function of GMCs/self-gravitating clouds (owing to a "pileup" of clouds at the scale of ~ 100 pc where the fragmentation cascade is artificially truncated), and numerically spurious stellar clusters (see also Nelson 2006; Lukat & Banerjee 2016). More dramatically, if we adopt $N \gtrsim 10$ (at the resolution shown here), then we are, by construction, forcing the gas to be smooth on scales larger than the (physical) gas disk scale height – in other words, we artificially suppress *any* substructure in the galactic disks, which in turn artificially suppresses star formation everywhere except at galaxy centers.

All of these effects weaken as we go to higher resolution (further indicating that they are spurious), but they vanish with no ill effects if we simply do not employ artificial pressure at all.

7 DO THE DETAILS OF COOLING & CHEMISTRY MATTER?

7.1 Cooling, Heating, and Self-Shielding Physics

We now consider how the *details* of cooling alter our predictions. Of course, it goes without saying that cooling is critical for galaxy and star formation – "turning off cooling" would produce no galaxies! Our interest here is in exploring whether the more subtle details of cooling, which are often uncertain at the factor ~ 2 level and treated differently or incompletely in different codes (for example, some include chemical networks for low-temperature and/or non-equilibrium cooling, some include only pre-tabulated cooling tables, some include additional sub-dominant cooling processes), ultimately have significant effects on our galaxy-scale predictions. Effects on observables such as line diagnostics can be more substantial, but are not our focus here (but see e.g. Richings & Schaye 2016).

Figs. 27-28 show the results of turning on or off different portions of the cooling physics in our simulations. Specifically, we compare our default FIRE-2 cooling physics described in § 2.3 to simulations adopting an extremely simplified toy model. In the toy model, we place all gas on a single, invariant cooling curve $\Lambda(T, Z, I_\nu, n) \rightarrow \Lambda(T)$ for fully ionized, non-self-shielded gas (so

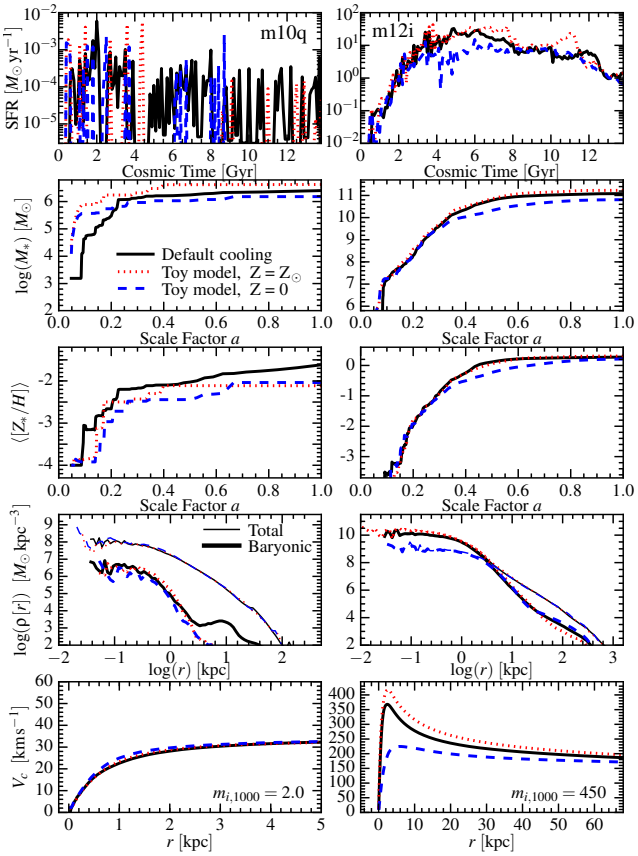

Figure 28. Effects of the details of cooling & chemistry (as Fig. 27), on cosmological simulations of a dwarf (**m10q**) and MW-mass (**m12i**) galaxy, as Fig. 4. We compare our default, full-physics & chemistry treatment, to the extremely simplified "toy" cooling model from Fig. 27 & § 7 (single, solar metallicity $Z = Z_\odot$ cooling curve, with no self-shielding so no low-temperature cooling or chemistry). We also compare a version of the toy model assuming $Z = 0$ (primordial abundances only, no metal-line cooling at low or high temperatures). The simplified cooling model produces remarkably similar galaxies to the full-physics model, because cooling times are always shorter than dynamical times so as long as some cooling channel exists, the details of cooling have little dynamical effect. The stellar masses are slightly higher with the $Z = Z_\odot$ toy model, owing to solar metallicity being assumed in all gas – this is higher than the metallicity of large-scale super-bubbles in **m10q** or outer parts of the extended "hot halo" of **m12i**, artificially enhancing cooling at $T \sim 10^5 - 10^7$ K. Conversely, with $Z = 0$ cooling only, SNe bubbles and the hot halo cool less efficiently; notably in **m12i** this reduces the SFR around $z \sim 1$, where the dense central bulge forms, flattening the rotation curve (more similar to a high-resolution simulation). Note that taking our "full cooling" model and *only* disabling high-temperature ($T > 10^4$ K) metal-line cooling produces nearly identical results here to the "toy" $Z = 0$ model.

there is essentially no cooling to $T \ll 10^4$ K), with a single metallicity $Z = Z_\odot$ or $Z = 0$ (primordial cooling only). We also remove the requirement that stars form in molecular/self-shielded gas, since this is not self-consistently computed in the toy model. This toy model is not intended to be physically correct! Rather, it is intended only to illustrate how even radical changes to cooling physics (much more extreme than typical model differences in cooling) have relatively small impacts on galaxy formation.

For dwarf galaxies, with either $Z = 0$ or $Z = Z_\odot$ (Fig. 28), we see there is essentially no effect on the galaxy formation history (since these are low-resolution, bursty galaxies, the differences in final mass or metallicity are essentially stochastic). For MW-

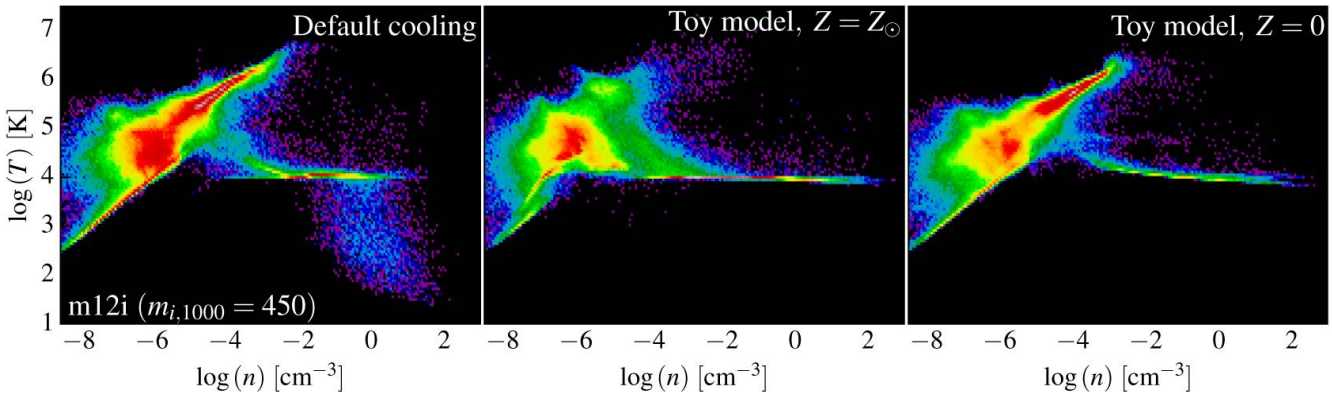

Figure 29. Temperature-density phase diagram of all gas in the simulation box at $z = 0$, for the **m12i** simulations in Fig. 28 with our default, full cooling physics implementation and the "toy model" which puts all gas on a single cooling curve, with no self-shielding, at a uniform metallicity $Z = Z_\odot$ or $Z = 0$. As expected, without self-shielding, the toy models produce no cold gas ($T \ll 10^4$ K). However, this is still "cold" as far as providing dynamical support to the disk (thermal $Q < 1$) is concerned, $t_{\rm cool} \ll t_{\rm dyn}$ remains true, and so even at $\sim 10^4$ K gas can and should physically fragment efficiently (isothermally) to Jeans masses $\sim 10^5\,M_\odot$ (less than the particle mass here) at the mean SF densities of the simulations. Therefore, this produces little or no dynamical effect in the simulations. The $Z = Z_\odot$ run noticeable suppresses the hot gas at $T \sim 10^6$ K and densities $n \sim 10^{-4} - 10^{-2}\,{\rm cm}^{-3}$ – much of this is gas in the extended CGM and nearby IGM (around $R_{\rm vir}$) which has been shocked (by SNe & O-star/AGB winds but also the halo virial shock), and is highly *sub-solar* in metallicity, so the toy model over-estimates the cooling rates by a factor $\sim 10-30$. The "primordial" ($Z = 0$) model, on the other hand, is closer to correct in the hot halo at $\sim R_{\rm vir}$, but significantly under-estimates the cooling rate for SNe bubbles in the gas disk and inner halo (while the shape of the warm/hot gas distribution is similar, there is much less gas in the warm phase compared to the Default cooling run).

mass galaxies, we find the same in both full cosmological simulations (Fig. 28) or re-starts near $z = 0$ with fixed initial conditions (Fig. 27), *if* we adopt approximately the "correct" metallicity in and around the galaxy, $Z = Z_\odot$. However for MW-mass systems, the primordial cooling-only ($Z = 0$) run does differ; this is discussed below. We have also examined the visual galaxy morphology, and wind outflow rates and covering factors, in each run and find similar results.

Fig. 29 shows the $z = 0$ temperature-density phase diagram of the **m12i** runs: clearly, details of the phase structure differ, as expected. With the toy model, there is no dense, cold gas ($T \ll 10^4$ K), by construction. For the toy model with $Z = Z_\odot$, the hot, denser gas has also cooled more efficiently – much of this was gas in the halo with lower metallicities which now radiates more rapidly than it should. But this does not appear to have a strong *dynamical* effect (there is some increase in the late-time SFR, stellar mass, and circular velocity evident in Fig. 28 in this run, owing to additional cooling, but the effect is weak).

This weak or negligible dependence of galaxy properties on the details of cooling physics (for approximately the correct metallicity) is consistent with previous studies. For example, both Hopkins et al. (2011) and Hu et al. (2016) added/removed photo-electric heating, and Hopkins et al. (2012c) added/removed molecular cooling at low temperatures, and all found this had essentially no effect on SFRs, wind outflow rates, and other galaxy properties. A more detailed study by Glover & Clark (2012) considering each microphysical cooling mechanism in turn showed that as long as *some* cooling channel exists, the details of that cooling are largely irrelevant in the dense ISM.

The reason for this is simple: for almost all gas (by mass) in galaxies, the cooling time is much shorter than the dynamical time ($t_{\rm cool} \ll t_{\rm dyn}$), and the gas is super-sonically turbulent. Under these conditions, the details of the cooling physics have essentially no effect on dynamics – since cooling is not a "rate-limiting" step, changing the rate by an order of magnitude in either direction does nothing, practically speaking. In fact, we find our predicted SFRs are essentially identical so long as the cooling model captures two key effects: (1) $t_{\rm cool} \lesssim t_{\rm dyn}$ for all gas with $T \ll 10^7\,{\rm K}\,(n/0.01\,{\rm cm}^{-3})$ (although for SNe blastwaves, the more relevant comparison is to the remnant expansion time, giving $T \lesssim 10^6$ K), and (2) gas in this rapidly-cooling regime has equilibrium temperatures $\lesssim 10^4$ K or colder (the colder gas has no significant thermal pressure support, so assigning it $T \sim 10^4$ does not appreciably change its large-scale dynamics).

Basically, in gas with temperatures $\lesssim 10^6 - 10^7$ K, details of the cooling physics are not important. Above this threshold, however, we do enter the regime where the cooling time becomes comparable to or longer than the dynamical time – then cooling can be important. This matters, potentially, for (1) super-bubbles forming and hot winds escaping galaxies, and (2) "hot halos" of virial shock-heated gas around galaxies. However, for winds (case (1)), so long as the cooling time is longer than the dynamical time, hot (thermal pressure-driven) winds generally cool adiabatically as they expand out of the galaxy and entrain more material (Chevalier & Gardner 1974); so again, the exact radiative cooling rate is not important (all that matters is that the transition from $t_{\rm cool} \ll t_{\rm dyn}$ to $t_{\rm cool} \gg t_{\rm dyn}$ happens at more or less the correct temperature and density). This is discussed in more detail in Paper II: metal-line cooling produces the dependence shown therein of SNe blastwave terminal momentum on metallicity, but this dependence is quite weak ($\propto Z^{-0.14}$), consistent with the argument above (also note that the low-resolution tests in Figs. 27-29 do not resolve the cooling radii of individual blastwaves, further minimizing the effects).

In hot halos (case (2)), the gas is quasi-hydrostatic, so the cooling rate in the halo center is the determinant of the rate at which new material is added to the galaxy. This is irrelevant in dwarf halos whose virial temperatures are $\ll 10^7$ K, but does become relevant in MW-mass halos. This is why we do see a significant effect in our **m12i** simulation if we adopt the toy model with $Z = 0$. We have re-run this simulation turning on and off each piece of the cooling physics in turn, and have verified that almost the entire change visible in Fig. 28 owes to the $Z = 0$ toy model having no high-temperature ($T > 10^4$ K) metal-line cooling. Interestingly, the change in the SFH from disabling metal-line cooling is nearly

identical to the results of our SPH comparison simulations (§ 5), or from going to much higher resolution (§ 4.1): the SFR is suppressed around $z \sim 1-2$, as the hot halo forms and in the "default" run, the SFR is dominated by recycled wind material from earlier star formation episodes. This wind material, if it is "trapped" in the inner CGM, is metal-rich so it cools rapidly and falls back into the galaxy, taking with it considerable "new" gas (with which it has mixed) as well. But in the toy model with $Z = 0$, there is no additional cooling from the metal-rich wind material, so this effect (incorrectly) never occurs. Similar conclusions have also been reached in previous studies (e.g. Choi & Nagamine 2009; Piontek & Steinmetz 2009; Schaye et al. 2010).

7.2 Effects of Numerical Metal Mixing

In § 2.6 we note that, in our "core physics only" simulations, using our default MFM hydrodynamic solver, passive scalars such as metals are "locked" to the gas resolution elements into which they are initially deposited (e.g. the nearest few elements to each SNe). This occurs because the MFM method is a finite-mass method, i.e. advective fluxes vanish *identically*. But since we have finite resolution, this fails to capture e.g. small-scale turbulent eddies and microphysical diffusion of metals between the boundaries of two neighboring (large) resolution elements. In finite-volume methods such as grid-based Eulerian codes, or moving-mesh codes, or the meshless finite-volume "MFV" method in GIZMO (which is nearly identical to MFM, differing only in how the face is assumed to deform when solving the Riemann problem between cells) there is an explicit advective (mass) flux, hence there is always some *unavoidable* mixing at the boundary (in fact, it is well known that such methods will tend to over-mix, relative to a converged solution, at any finite resolution).

Under some circumstances, it may be desirable to explicitly model this microphysical metal transport between resolution elements. In Appendix F3 we describe how this is implemented as an "additional physics" option in FIRE-2. Essentially this follows the Smagorinsky (1963) approximation, where we explicitly solve a diffusion equation between cells with the diffusivity $\kappa_{\rm turb}$ set to the resolution-scale "eddy diffusivity" ($\sim \ell\,|\Delta v_{\rm turb}(\ell)|$; i.e. assuming the mixing time scales with the eddy turnover time, based on the local spatial resolution and velocity field). This is tested directly in GIZMO with our same MFM solver in idealized, converged "turbulent box" simulations in Colbrook et al. (2016) and Rennehan et al. (in prep.). An upper limit to the mixing between cells is enforced, equal to what would be obtained by the GIZMO MFV solver for the same cells and same timestep – in other words, the *maximum possible* diffusivity of the added term in MFM is simply the un-avoidable diffusivity inherent to the MFV method (itself very similar to that in popular moving-mesh codes). This allows us to study how this particular numerical diffusion term, which is always present in hydrodynamic methods with explicit mass fluxes, alters our predictions.

Fig. 30 shows the effects of adding this term, in several of our simulations at both dwarf and MW mass scales, on the galaxy formation history and $z = 0$ mass profiles. Fig. 31 shows the effects on the $z = 0$ ISM and CGM gas phase distribution and visual morphologies in a MW-mass galaxy (there is no detectable difference in dwarfs, consistent with other properties we have surveyed). We see essentially no systematic effect in any gross galaxy property (star formation history, stellar mass, mass profile, rotation curves, visual morphology of gas or stars), including the mean metallicity itself (both in gas and stars). This is true at all mass scales we have explored and at all resolution levels. There is *slightly* more cool gas at $z = 0$ in the disk of the **m12f** simulation with metal diffusion, which could owe to metal mixing promoting additional cooling from the hot halo, but we caution that this effect is similar in magnitude to stochastic variations between runs. It is worth noting that the cold gas within the disk ($T \lesssim 10^4\,{\rm K}$, $n_H \gtrsim 0.1\ {\rm cm^{-3}}$) forms a slightly tighter (more visually obvious) cooling sequence; this is because the metal diffusion smooths local particle-to-particle variations in gas-phase metallicity in e.g. the same GMCs, and the metals are the dominant coolant in this regime.

Su et al. (2016) explore in greater detail the effects of this metal-mixing term in both isolated (non-cosmological) disks and cosmological simulations with the FIRE-2 physics. They specifically compare the bivariate distribution of gas temperatures and densities in the ISM (at several redshifts); galactic wind outflow rates, phases, density distributions, and velocity distributions; turbulent velocity distribution functions in the ISM; and star formation histories. They conclude in all cases that the metal-mixing terms have weak or negligible effects. In future work, Escala et al. (in prep.) will present the actual abundance ratio *distribution functions* within the galaxies here; there, it is clear that the mixing terms do have some effect (as one might expect). However because the mixing terms conserve (by construction) the mean metallicity and total metal mass, and we have already shown that the metallicity enters only relatively weakly into dynamical effects via its effect on the cooling rates (see § 7), the dynamical effect of these mixing terms on any other galaxy properties is weak.

Fig. 32 further demonstrates this by considering variations to the standard metal-mixing term, at both dwarf and MW mass scales. Specifically we consider (in addition to the "core physics only" and "metal diffusion" runs) cases where we add the metal diffusion term but arbitrarily multiply the diffusivity $\kappa_{\rm turb}$ by an additional factor of 10 or 30. We also consider a case where we replace the standard adaptive eddy diffusivity (defined in Appendix F3) with a constant $\kappa_{\rm turb} = 10^{24}\,{\rm cm^2\,s^{-1}}$ (i.e. simply treat the metals as if they obey a strict, constant-diffusivity diffusion equation); we also remove the MFV-based limiter above. This value of $\kappa_{\rm turb}$ corresponds crudely to typical values of the turbulent diffusivity estimated by Eq. F5 in Appendix F3 on the maximal turbulent scales (for $\ell \sim 100\,{\rm pc}$ and $|\Delta v_{\rm turb}(\ell)| \sim 10\,{\rm km\,s^{-1}}$). Finally, we consider a case where we include no diffusion equation, but instead we solve the metal flux *exactly* as we would if we used the MFV hydrodynamic method (with explicit mass fluxes), and include the resulting metal fluxes (even though all *other* fluxes use the MFM solution, identical to our "standard" runs). In other words we always transfer the metals that would be "numerically diffused" in a finite-volume code. In each of these cases, we see no effect on the properties in Fig. 32.

7.3 Yields and Explicit Abundance-Ratio Tracking

As noted above, we track 11 species on-the-fly with yield tables for SNe (Ia & II) and stellar mass-loss given in Appendix A. But nucleosynthetic yields (even IMF-averaged) have very large uncertainties, especially at progenitor metallicities far from solar. For example, the Woosley & Weaver (1995) yields differ from those in Nomoto et al. (2006) even at $Z = Z_\odot$ by $\sim 0.4\,{\rm dex}$ for Mg and Ne. However, given the very weak dependence of our results on the details of the cooling curve shape (especially at low temperatures; § 7.1) and detailed metal mixing (§ 7.2), we do not expect this to have large dynamical effects on other galaxy properties. Moreover, even high-temperature metal-line cooling (which is important to include) is dominated by O, which constitutes most of the metal mass and therefore is better constrained (differing by $< 10\%$ be-

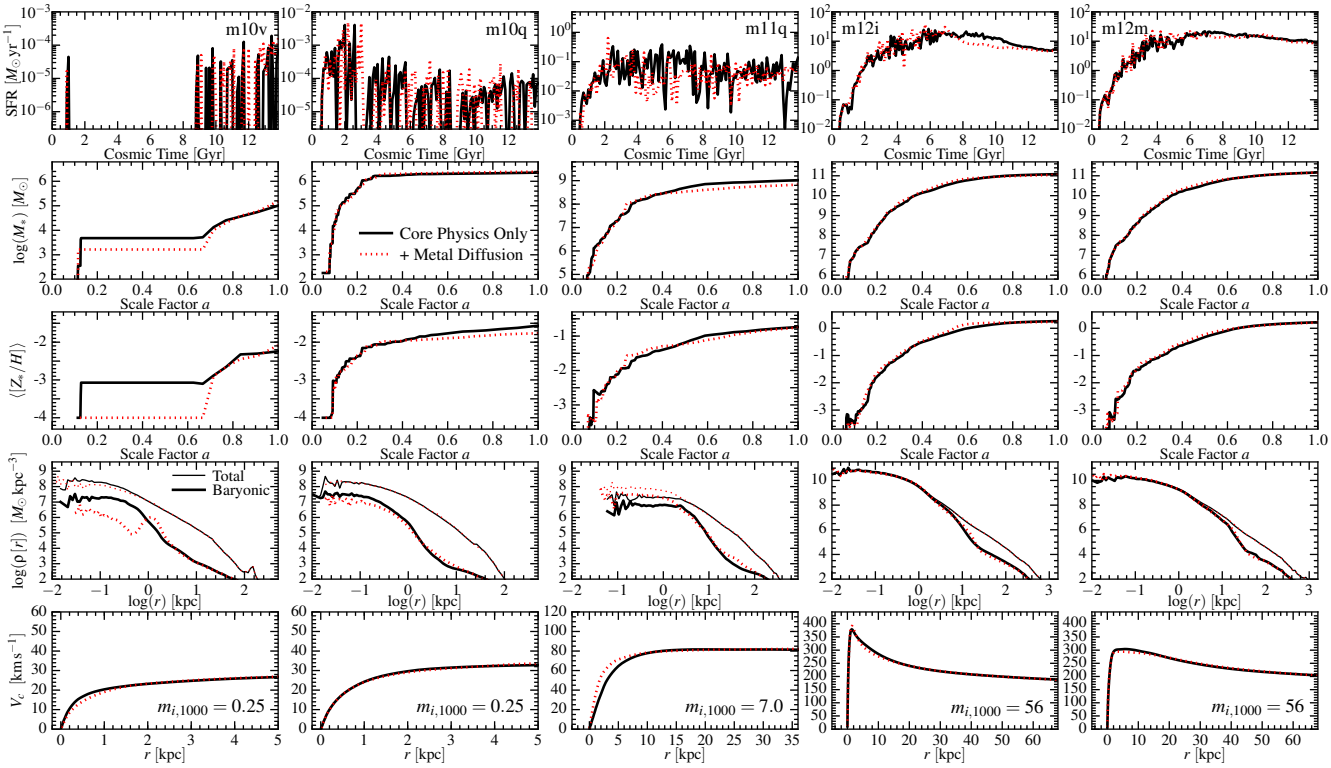

Figure 30. Comparison of predicted galaxy formation histories as Fig. 4, with and without an explicit numerical "metal diffusion" term. Our "core physics only" runs using the MFM hydrodynamic solver follow fixed-mass elements (so there is no advection between elements and metals are strictly "locked" to the element they were injected into via stellar mass-loss). The "metal diffusion" runs add an explicit passive-scalar diffusion term to account for sub-resolution diffusion and turbulent mixing of metals between neighboring elements (see Appendix F3). This is flux-limited so that the diffusivity is never larger than would be obtained from simply numerical diffusivity in our MFV (moving-mesh or AMR-like) hydrodynamic solver. There is no systematic effect at any mass scale. In addition to the runs shown we have also compared **m09** and **m12b,c,f,q** and find the same. Su et al. (2016) examine non-cosmological disks and our **m10q** and **m12i** runs and show the same, also for the gas and stellar morphology, ISM density distribution in different phases, galactic wind mass-loadings and velocity distributions, and ISM turbulent velocities.

tween the Woosley & Weaver 1995 and Nomoto et al. 2006 models, for example). We have verified this directly by re-running our **m10q** ($m_{i,1000} = 0.25$) and **m12i** ($m_{i,1000} = 56$) simulations, replacing the default Nomoto et al. (2006) yields for core-collapse SNe with those from Woosley & Weaver (1995). In all properties surveyed here, we see no detectable difference.

Given this and our comparisons in § 7.1, it appears that as long as the total metallicity is approximately correct, differences in the detailed abundance ratios introduce no major dynamical effects. We have therefore re-run the same simulations, ignoring all detailed abundance information and simply tracking the total metal abundance, then assuming solar abundance ratios in the cooling and other relevant routines. We find this produces nearly identical results to our default simulations (in which the cooling is explicitly solved species-by-species).

To leading order, then, detailed abundance patterns are important as *tracers* for predictions of various observables, but do not introduce important dynamical effects or uncertainties into predictions of other galaxy properties.

8 EFFECTS OF THE STAR FORMATION ALGORITHM

It has been extensively demonstrated in the literature that, provided the Toomre scale is resolved and stellar feedback is treated explicitly (as in our simulations), the exact resolution-scale SF model has essentially no effect on predicted galaxy-scale SFRs or other quantities. For the sake of completeness, we demonstrate this here, but refer readers to Hopkins et al. (2011, 2013d,e) and Kim et al. (2013a) for more extensive discussions (see also § 4.1.3).

Figs. 33-34 show the effects of varying the resolution-scale star formation model in the simulations, in both full cosmological runs at different mass scales and resolution, and re-starts of our MW-mass simulation at late times (guaranteeing an identical initial condition for the comparison). We compare our default model from § 2.4 (requiring gas be self-gravitating, self-shielding/molecular, Jeans unstable, and have density $n > n_{\rm crit}$ with $n_{\rm crit} = 1000\,{\rm cm^{-3}}$), to variations with $n_{\rm crit} = 5 - 1000\,{\rm cm^{-3}}$; turning on/off the self-gravity (virial), molecular, and Jeans criteria; and arbitrarily multiplying the SFR per free-fall time $\dot{\rho}_*$ in the gas which meets these criteria by a constant factor $\epsilon_{\rm sf} = 0.01 - 100$.

In every case, the predicted SFR and all other galaxy properties we examine here are essentially identical. This has now been shown repeatedly and in many different contexts, with simulations using different detailed implementations of star formation and stellar feedback; mass resolution ranging from sub-solar to $\sim 10^6\,M_\odot$; isolated (non-cosmological), galaxy-merger, and fully-cosmological simulations; circum-nuclear simulations of star formation around AGN; and simulations using very different feedback mechanisms (removing SNe, stellar mass-loss, or HII photoheating); and in different codes (see e.g. Saitoh et al. 2008; Shetty & Ostriker 2008; Hopkins et al. 2012b, 2013c, 2011, 2013a,d,e, 2016; Federrath & Klessen 2012; Agertz et al. 2013; Kim et al. 2011, 2013a; Kim & Ostriker 2015). We specifically showed the

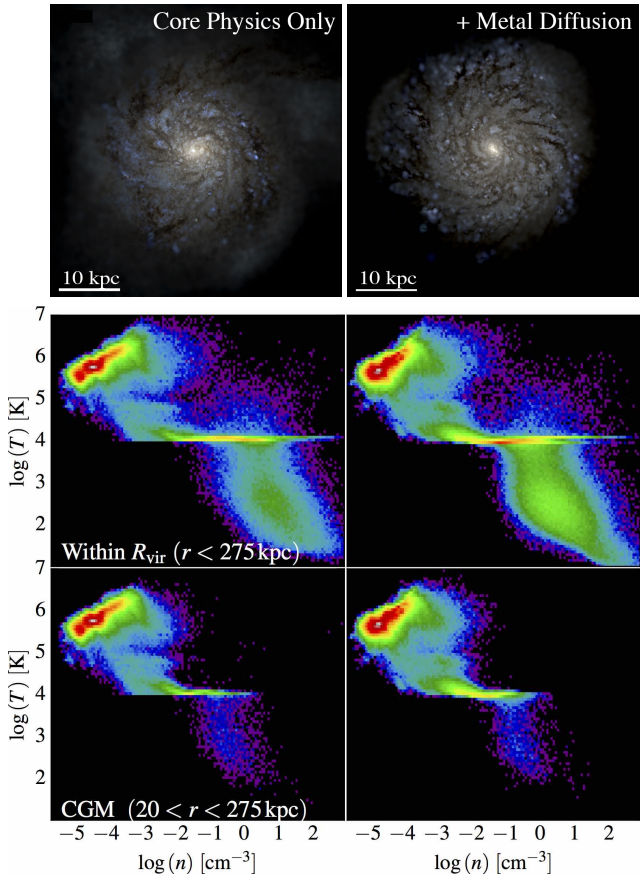

Figure 31. *Top:* Visual morphology (as Fig. 9), at $z = 0$, for the **m12f** ($m_{i,1000} = 56$) simulations run with our "core physics only" (*left*) or core physics with the additional metal diffusion term as Fig. 30 (*right*). *Bottom:* Phase diagram of the ISM & CGM (as Fig. 26) for the same. The diagrams are qualitatively very similar but there is slightly more cool gas in the ISM at $z = 0$ in the runs with metal diffusion, owing to its allowing marginally enhanced cooling from the hot halo that forms at late times in massive galaxies. There is no detectable difference at dwarf masses.

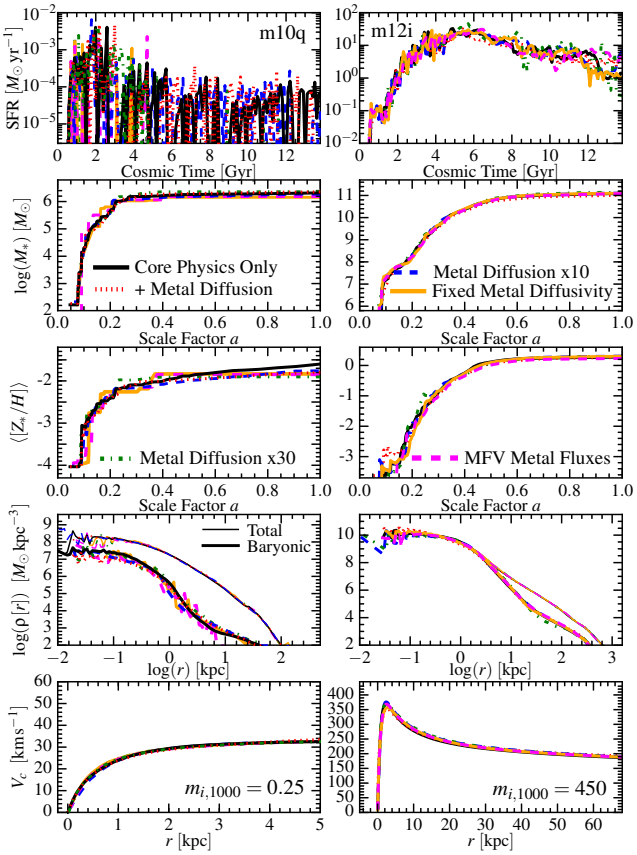

Figure 32. Comparison of the effects of explicit numerical metal diffusion, as Fig. 30, at dwarf (**m10q**) and MW (**m12i**) mass scales, as a function of the diffusivity. We compare: (1) *Core Physics Only*: metals strictly locked to original injection element. (2) *Metal Diffusion*: the default numerical implementation in Appendix F3. (3) *Metal Diffusion x10*: as "Metal Diffusion" but with the numerical diffusion coefficient increased by a factor 10. (4) *Metal Diffusion x30*: as "Metal Diffusion" but with the numerical diffusion coefficient increased by a factor 30. (5) *Fixed Metal Diffusivity*: replacing the adaptive numerical diffusion prescription with a pure isotropic diffusion equation of fixed diffusivity $\kappa_{\rm turb} = 10^{24}\,{\rm cm^2\,s^{-1}}$, comparable to the large-scale eddy diffusivity for ISM turbulence with $\sigma \sim 10\,{\rm km\,s^{-1}}$ and driving scale $\sim 100\,$pc. (6) *MFV Metal Fluxes*: explicitly adding the cell-to-cell metal advection fluxes that *would* appear if we solved the hydrodynamic equations with GIZMO's "meshless finite volume" (MFV) method, instead of our default MFM method (which has vanishing advective fluxes). This is approximately the "numerical metal diffusion" that would un-avoidably appear in a moving-mesh type code. In all cases, the effects are weak.

same was true in our FIRE-1 simulations in Hopkins et al. (2014). This is also consistent with simulations of individual GMCs (which track star formation on much smaller scales), and analytic "multi-free-fall" models of star formation, in which independent clumps collapse on their own local free-fall times (see e.g. Federrath & Klessen 2012, 2013; Faucher-Giguère et al. 2013; Guszejnov & Hopkins 2016; Grudić et al. 2016).

This is because the rate-limiting step in star formation does not occur at the resolution limit of the simulations; rather it is the formation of the first dense, self-gravitating objects (large GMC complexes, which have dynamical times $\sim 100\,$Myr). Because these are *resolved*, and our force softening is fully-adaptive (and we do not force some artificial lower-limit to the softening), a self-gravitating cloud or sub-clump will continue to collapse, as it should, on its local free-fall time, until it either forms stars or is destroyed by feedback. This means (absent star formation or sufficient feedback) the densities in such a clump will (correctly) become arbitrarily high and the internal dynamical times will become arbitrarily short, within a finite physical time (the original parent cloud free-fall time). So whatever value of $n_{\rm crit}$ we set will eventually be exceeded, and even if $\epsilon_{\rm sf} \ll 1$ (so the cloud collapses faster than it forms stars), the actual physical SF timescale will become arbitrarily short within the clump as it collapses. Provided star formation *can* occur, then the simulations above have (without exception) shown that feedback self-regulates the level of star formation on the galactic scale. Stars form until sufficient numbers of young stars are present that their feedback destroys the parent clouds and prevents new collapse (locally "lifts up" the cloud or molecular disk). This self-regulating behavior depends on the strength of feedback – we will show below that changing the strength of feedback immediately changes the equilibrium SFR in a galactic disk – but is independent of how individual stars form in dense gas.

Of course, the ratio of SFR to dense gas near the resolution limit depends, by construction, on the SF model; comparison to dense gas tracers within galaxies (e.g. HCN as opposed to CO) in previous work appears to favor the "default" normalization and SF criteria we adopt here (see the comparison in Hopkins et al. 2013e). But this is only relevant within the densest resolved gas clumps.

From the arguments above, the key criteria for the robust be-

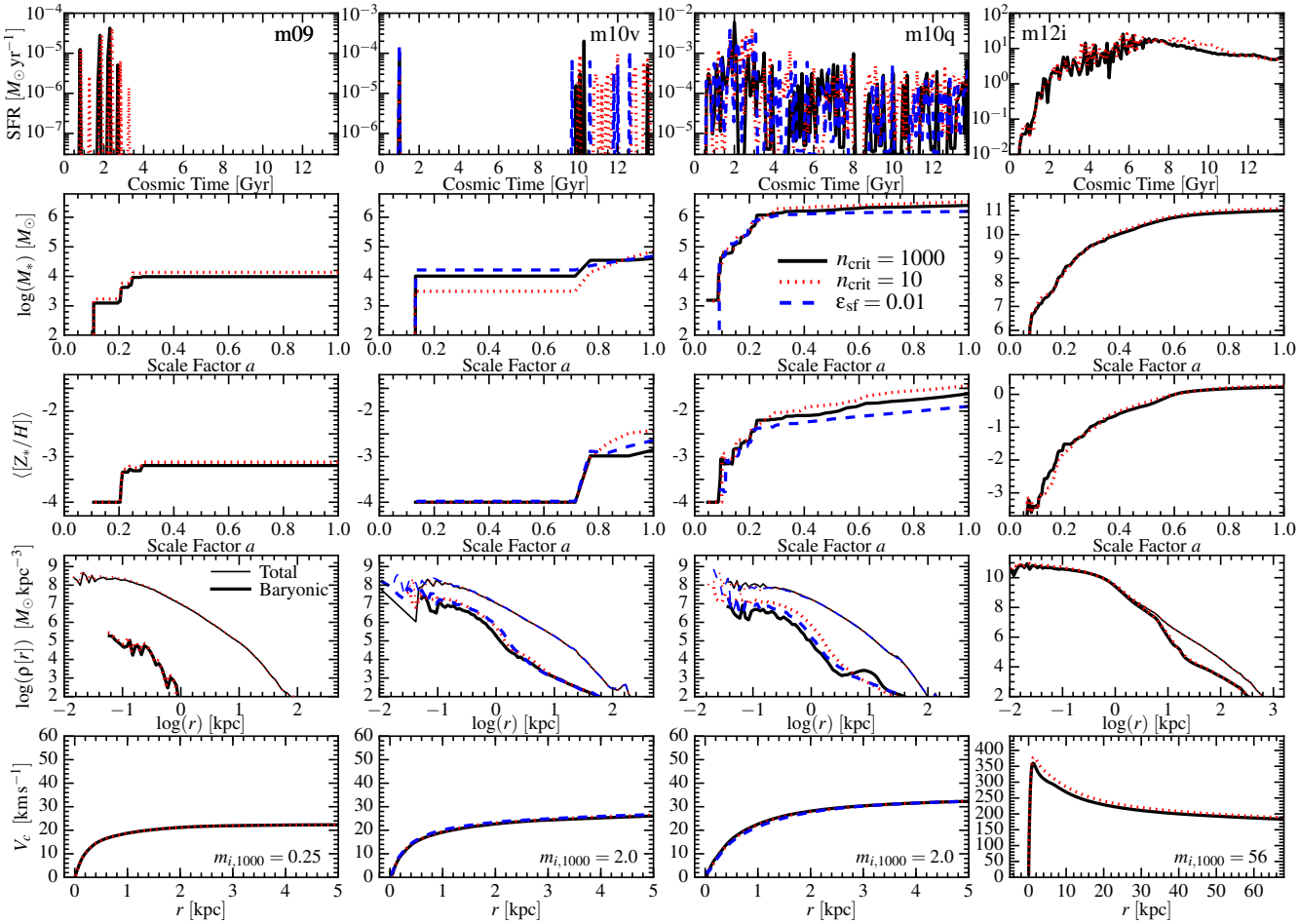

Figure 33. Comparison of galaxy properties in cosmological FIRE-2 simulations (as Fig. 4) as a function of the resolution-scale assumptions for individual star particle formation. We compare **m09**, **m10v**, **m10q**, and **m12i** (Table 1). For each, we compare a run with a minimum density for star formation of $n_{crit} = 10$ or $n_{crit} = 1000\,cm^{-3}$; recall, in all cases the SF is still only allowed in self-gravitating, self-shielding gas. We have repeated this experiment for **m10q** and **m12i** at lower resolution, and for **m11q**, **m11v**, **m10_halo1146** with $n_{crit} = 100$ and $= 1000$, and reach the same conclusion in every case. For **m10v** and **m10q**, we also compare a simulation in which we arbitrarily multiply the SFR in the "eligible" gas (sufficiently dense, self-shielding/molecular, locally self-gravitating) by a factor $\epsilon_{SF} = 0.01$ (i.e. artificially "slow down" the SF in the collapsing gas by a factor of ~ 100, relative to its dynamical time). Consistent with extensive studies in the FIRE-1 and previous simulations, these choices have no effect on our predictions.

havior we see are: (1) the threshold for SF (n_{crit}) be higher than the mean disk density, *or* a self-gravity criterion like what we use is adopted, so the identified regions are likely to be actually collapsing; (2) the dominant (largest, Toomre) scales of fragmentation are *resolved* and self-gravity is *allowed* to collapse the gas to high densities (i.e. no artificial, large gravitational softening is imposed; see § 4.2 above); and (3) stellar feedback is treated explicitly.

9 STELLAR FEEDBACK

Many previous studies have now shown that feedback is the most important determinant of star formation in galaxies (see references in § 1). In this section we therefore explore how basic feedback physics changes our predictions; in Paper II-Paper III, we examine more subtle physical and numerical aspects of the *implementation* of said feedback.

9.1 Feedback "Strength"

In Fig. 35, we simply re-start our **m12i** simulation near $z = 0$ as Fig. 10, but arbitrarily multiply *all* feedback rates (e.g. SNe rates, wind kinetic luminosities and mass loss rates, stellar luminosities at all frequencies) by a factor of ~ 3 ("strong") or $\sim 1/3$ ("weak"), relative to the "default" values (which, recall, are taken directly from stellar evolution models without adjustment). We stress that there is no physical motivation for this – this is a larger shift than most uncertainties in stellar evolution – but we consider the counter-factual case purely for illustrative purposes. Exactly as expected in "feedback-regulated" scenarios, this produces a corresponding direct shift in the SFR – the SFR is self-regulating at the level where feedback offsets gravitational collapse, so if feedback is $3\times$ stronger "per star," then $3\times$ fewer stars form when the equilibrium is realized. This is almost exactly what we see (the shift in SFR is very slightly sub-linear, owing to non-linear effects). This is consistent with many previous simulation and analytic studies (see Hopkins et al. 2011, 2012c; Faucher-Giguère et al. 2013; Shetty & Ostriker 2008; Kim et al. 2011, and references therein).

9.2 The Role of Different Feedback Mechanisms: Galactic Winds & Cosmological Timescales

In Fig. 36 we consider how turning off each feedback mechanism in turn alters galaxy evolution in fully cosmological simulations. In both dwarf and MW-mass galaxies, we clearly see SNe have the most dramatic effect. This is because what ultimately regulates

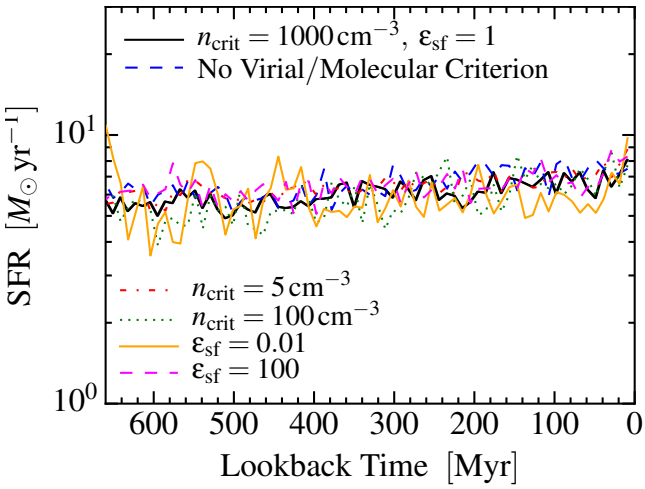

Figure 34. Effects of the resolution-scale assumptions for individual star formation, in re-starts of a MW-mass galaxy (**m12i**) at late times, as Fig. 10, keeping fixed mass resolution $m_{i,\,1000} = 56$ but changing the SF algorithm. In our "default" model gas which is self-gravitating (sub-virial), self-shielding/molecular, dense ($n > n_{\rm crit} = 1000\,{\rm cm}^{-3}$), and Jeans unstable forms stars at a rate $\dot{\rho}_* = \rho_{\rm mol}/t_{\rm ff}$. We compare: (1) removing the self-gravity & molecular restrictions (all dense gas can form stars), (2-3) lowering/raising $n_{\rm crit} = 5 - 1000\,{\rm cm}^{-3}$, and (4-5) multiplying $\dot{\rho}_*$ by an arbitrary factor $\epsilon_{\rm sf} = 0.01 - 100$. As in Fig. 33 and our previous studies, this has no effect on the galaxy-scale SFR because it is feedback-regulated.

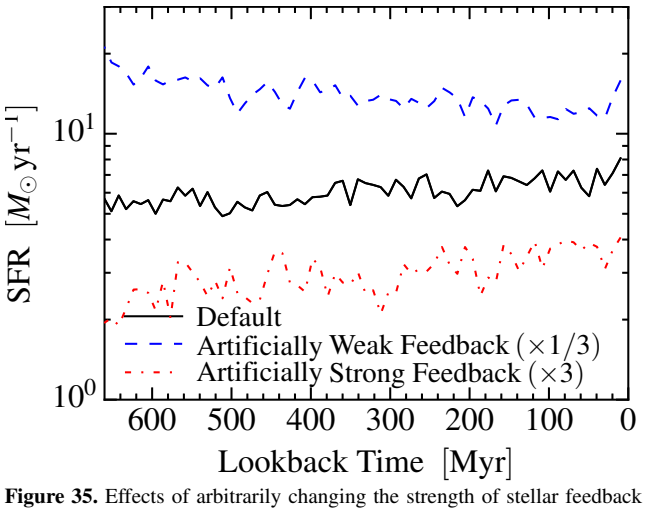

Figure 35. Effects of arbitrarily changing the strength of stellar feedback on a re-start of an identical **m12i** simulation (with $m_{i,\,1000} = 56$) at low redshift (i.e. guaranteeing identical initial galaxy properties), as Fig. 10. Here, we compare our "default" model, in which all feedback physics, rates and energetics are taken *directly* from stellar evolution models without re-adjustment, and compare it to two models where we arbitrarily multiply *all* feedback rates (e.g. SNe rates, wind kinetic luminosities and mass loss rates, stellar luminosities at all frequencies) by a factor of ~ 3 ("strong") or $\sim 1/3$ ("weak"). These are much larger than actual physical uncertainties in these quantities, we simply show it for illustrative purposes. Clearly SF is instantaneously feedback-regulated: the steady-state SFR is inversely proportional to the feedback strength, as expected.

galactic star formation efficiencies over a Hubble time is the competition between inflow and galactic outflows, and the high-speed outflows that are not simply recycled quickly are predominantly SNe-driven (Davé et al. 2011a). Without SNe, galaxies form far too many stars (factor > 100 in our **m10q** dwarf; factor ~ 2 in the MW-mass system, going from $\sim 30\%$ to $\sim 60\%$ baryon-to-star "conversion" efficiency); the stars form too early (everything collapses rapidly into early dwarf halos, and has mostly turned into stars by redshift $z \sim 3-4$); the resulting galaxy is too dense (rotation curves peak strongly, even in small dwarfs, and are never "gently rising"); and the metallicities are essentially constant across stellar mass because galaxies have similar formation histories and baryon conversion efficiencies (late-time metallicities are lower for massive galaxies because there is no SNe enrichment, only RSG/AGB enrichment). Likewise properties of the ISM (gas phases) and CGM (covering factors of outflows and metals, especially) are grossly discrepant with observations. We also note that because all of our galaxies are star-forming, the SNe Type II rate dominates over the Ia rate at all times. All of these results are consistent with many previous studies (see references in § 1). We discuss the effects of SNe in much greater detail in Paper II.

Turning off continuous stellar mass-loss (OB/AGB winds), the most obvious effect is that the late-time SFRs of both dwarfs and MW-mass galaxies are *suppressed*, relative to their "default" values. Despite the fact that OB winds carry (roughly) comparable energy and momentum flux to SNe, and AGB winds can do the same if the relative star-gas velocity is sufficiently large (see Appendix D), their *net* effect seems to be increasing the supply of gas mass within the galaxy which can eventually cool and form stars at late times. This is also consistent with previous studies, many of which have noted that OB winds tend to not be sufficiently well-confined to build up comparable momentum (after coupling) to SNe (Jungwiert et al. 2001; Ciotti & Ostriker 2001; Ciotti et al. 2009; Schaye et al. 2010; Leitner & Kravtsov 2011; Novak et al. 2011; Gan et al. 2014; Choi et al. 2015).

Removing all radiative feedback – including the effects of the meta-galactic UV background – has a dramatic effect on dwarf stellar masses, even at halo masses $\gtrsim 10^{10}\,M_\odot$ with $V_{\rm max} \gtrsim 40\,{\rm km\,s}^{-1}$ that should be well above the threshold for UVB "quenching" star formation. This will be explored in more detail in future work (Wheeler et al., in prep.). In short, the stellar masses increase by a factor ~ 10 (consistent with our FIRE-1 results; Hopkins et al. 2014). Unlike removing SNe, the qualitative shape of the SFH is similar to the "full feedback" case – the stars do not form uniformly early, and the SFH is still "bursty," but the bursts reach much larger amplitude, owing to the much larger gas supply. Removing only *local* radiative feedback (keeping the UVB, but removing radiative feedback from explicit star particles) has more subtle effects; we discuss these in detail in Paper III. The stellar masses do not change as much, but in the dwarfs, the SFR truncates at early times and the resulting $z = 0$ metallicity is suppressed. In MW-mass halos, the UVB has a much weaker effect. However, removing *local* radiative feedback produces a clear, significant increase in the central V_c – the galaxy center is exactly the region where we expect massive, dense GMCs which must be pre-processed by radiative feedback in order to reach low densities where subsequent SNe can efficiently expel the material (Thompson et al. 2005; Murray et al. 2005; Hopkins et al. 2011). The dependence is similar in FIRE-1, although there we saw a stronger dependence of stellar mass on radiative feedback; the difference owes to the improvements in the treatment of SNe which allow, in massive GMCs, for the first SNe to play a similar role to radiative feedback and "pre-process" the cloud for subsequent SNe.

As Fig. 36 clearly demonstrates, our conclusions about the systematic effects of different feedback mechanisms are independent of the hydrodynamic method (MFM vs. SPH), even where the

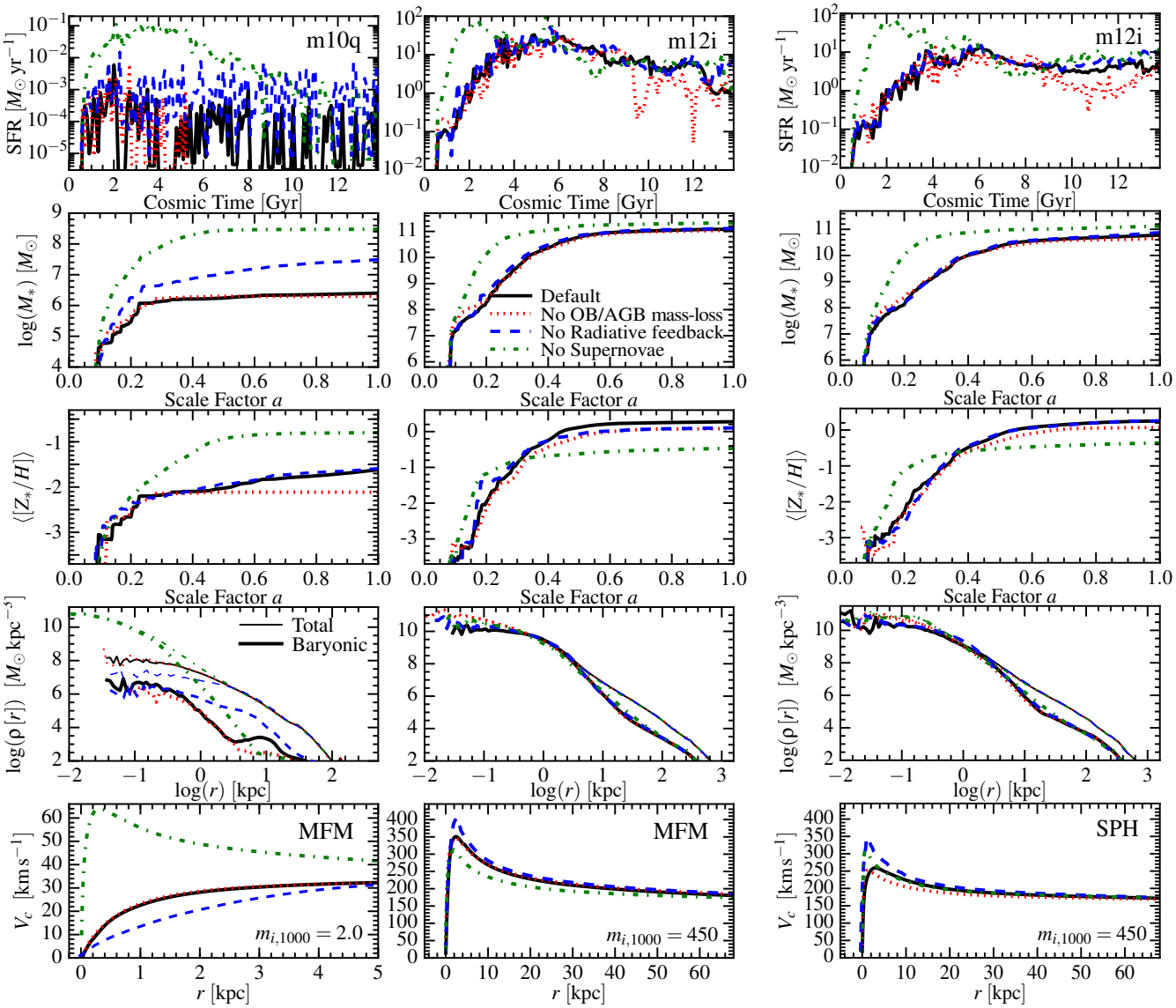

Figure 36. Effects of different stellar feedback physics on cosmological galaxy formation histories, as Fig. 4. We compare a dwarf (**m10q**) and MW-mass galaxy (**m12i**), both run with our default, higher-order accurate hydrodynamic solver (MFM; *left*); but we also show a set of simulations of **m12i** with the same physics variations using the smoothed-particle hydrodynamics (SPH; *right*) method, to demonstrate that even though the hydrodynamic solvers produce differences in massive galaxies, the qualitative effects of different feedback mechanisms are identical (independent of the hydrodynamic method). **(1)** *Default:* Our default models include all standard stellar evolution processes. **(2)** *No OB/AGB mass-loss:* Removing continuous stellar mass loss (both OB and AGB-star winds) produces slightly lower metallicities (owing to the lack of recycling) and significantly lower late-time SFRs – it appears the primary role of stellar mass-loss is to provide an additional source of gas fueling late-time SF in both dwarfs and MW-mass systems. **(3)** *No Radiative feedback:* This removes all radiative feedback (radiation pressure as well as photo-ionization and photo-electric heating by local particles and the meta-galactic background). In dwarfs (even with $V_{\rm max} \sim 40\,{\rm km\,s^{-1}}$, shown here), removing the photo-ionization heating (dominated by the UVB) produces $\sim 10\times$ larger SFRs and stellar masses (producing large bursts that make a core and lower V_c). In massive galaxies, the effects are weaker but removing radiation pressure produces significantly *higher* central densities (more strongly-peaked rotation curves in the central $\sim 5\,$kpc). **(4)** *No Supernovae:* SNe clearly dominate on cosmological scales, as removing them produces orders-of-magnitude higher SFRs at early times, giving rise to runaway collapse to extremely high densities until the gas is depleted.

methods themselves produce significant differences. Likewise, we have checked the comparisons above are independent of resolution, star formation prescription, and details of the cooling functions.

9.3 The Role of Different Feedback Mechanisms: Self-Regulation Within Galaxies and the Kennicutt-Schmidt Relation

In Fig. 37 we repeat the experiment from Fig. 36 turning on and off different feedback mechanisms, but in restarts of the $z \sim 0$ MW-mass galaxy as Fig. 10. This allows us to separate non-linear, long-timescale cosmological effects on the SFR (e.g. galactic wind generation and recycling), from the *instantaneous* self-regulation of star formation *within* a galaxy.

As expected, removing SNe produces a systematically higher SFR. Removing OB/AGB winds leads, initially, to higher SFRs, as the additional feedback (gas heating via shocked winds) is no longer present (but the gas mass of the disk is still essentially fixed at its initial value); however the SFR then declines as the missing mass-loss is unable to "re-supply" gas lost to star formation. Most interesting, without radiative feedback, the SFR systemati-

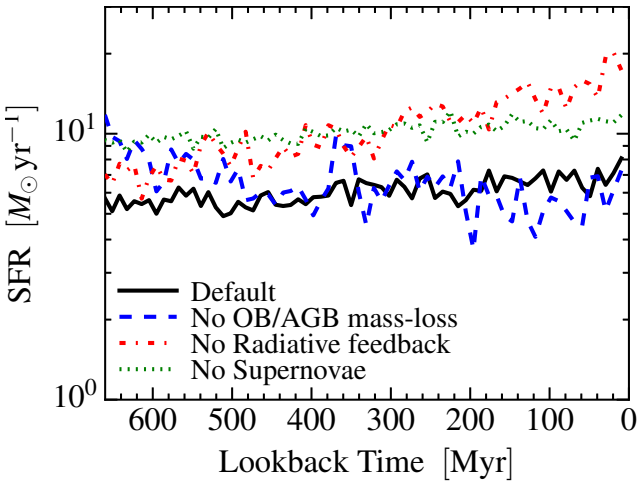

Figure 37. Effects of different stellar feedback physics on a re-start of an identical MW-mass simulation (**m12i**; as Fig. 35); we disable different mechanisms in turn as in Fig. 36. **(1)** *Default:* All physics included. **(2)** *No OB/AGB mass-loss:* Since the simulations start from identical initial conditions (identical gas supply), removing stellar mass-loss initially leads to a higher SFR (the OB winds cannot act as a source of feedback) – but by $\sim 500\,$Myr later, the net effect of removing the winds is to slightly decrease the SFR, as the winds cannot act continuously over this time to supply new gas mass to the disk. **(3)** *No (Local) Radiative feedback:* Here we still include a UV background, but remove only local (star particle) radiative feedback. We see a steady rise in SFR until this is forming more stars than the no-SNe case! Two effects occur: (1) GMCs are no longer efficiently disrupted by radiative feedback before SNe explode, therefore the SFR needed to self-regulate within a disk of a given surface density rises (see e.g. Hopkins et al. 2011), and (2) the outer atomic disk, which was stabilized by photo-heating from young stars in the galaxy, can now cool efficiently and form stars. **(4)** *No Supernovae:* The SFR is systematically larger, as expected, but by a relatively modest factor ~ 2. This is because other feedback mechanisms instantaneously can regulate the SFR of gas in the galactic disk. However, removing SNe eliminates almost all galactic winds – on cosmological timescales, this means the disk will increase in gas mass and the SFR will run away, as we saw in Fig. 36. Critically, this demonstrates that how feedback regulates SF *instantaneously within the disk* (e.g. the Kennicutt-Schmidt relation) is *not* the same as how feedback regulates SF *integrated over cosmological inflow/outflow/growth timescales* (e.g. the stellar mass-halo mass relation).

cally rises with time, comparable to or even exceeding the SFR of the no-SNe test! This is consistent with a number of studies of isolated (non-cosmological) galaxy simulations (see Hopkins et al. 2011, 2012c; Kim et al. 2013a; Agertz et al. 2013; Benincasa et al. 2016).

Why does the SFR appear more sensitive to radiative feedback in Fig. 37 and these isolated galaxy simulations, as opposed to Fig. 36? On cosmological spatial and temporal scales, a simple steady-state argument (e.g., Bouché et al. 2010; Davé et al. 2012; Feldmann 2013; Lilly et al. 2013; Feldmann 2015; Mitra et al. 2017) implies the sum of the time-average galactic SFR $\langle \dot{M}_* \rangle$ and superwind mass-loss rates $\langle \dot{M}_{\rm wind} \rangle \equiv \eta \langle \dot{M}_* \rangle$ (where η is the wind mass-loading, which can depend on arbitrary galaxy properties) should equal the inflow rate $\dot{M}_{\rm in}$ (we assume gas mass in the disk is steady-state; see Davé et al. 2011a). Thus $\langle \dot{M}_* \rangle \sim (1+\eta)^{-1} \dot{M}_{\rm in}$. However, on scales *within* the galactic disk (and timescales of order the galaxy dynamical time, much shorter than the Hubble time), the SFR is set by self-regulation via feedback, at the point where the momentum flux per unit area from feedback, $\sim (p_*/m_*)\,\dot{\Sigma}_*$ (where $p_*/m_* \sim 3000\,{\rm km\,s^{-1}}$ is the time and IMF-averaged flux of momentum *into the dense, star forming gas* per unit stellar mass formed) offsets the gravitational force per unit area $\sim G\,\Sigma_{\rm disk}\,\Sigma_{\rm gas}$ (an identical scaling is obtained assuming feedback must destroy star-forming clouds, and/or offset turbulent dissipation, in a $Q \sim 1$ disk; see Shetty & Ostriker 2008; Kim et al. 2011; Faucher-Giguère et al. 2013; Hayward & Hopkins 2017). Thus $\langle \dot{M}_* \rangle \sim (p_*/m_*)^{-1}\,G\,\Sigma_{\rm disk}\,\langle M_{\rm gas} \rangle$. Of course, these two scalings must be consistent: if the "instantaneous" $\dot{M}_*$ is "too low" compared to the cosmological scaling above, then $\dot{M}_{\rm in}$ is not offset by star formation+outflow and gas "piles up," raising $M_{\rm gas}$ until the SFR "catches up." If the radiative feedback contributes significantly to p_*/m_* (either directly or via pre-processing GMCs before SNe), then removing it lowers p_*/m_* and in turn increases the SFR given *fixed* $\Sigma_{\rm disk}$ and $M_{\rm gas}$, as in Fig. 37. However, if η is less sensitive to radiative feedback, the winds from higher $\dot{M}_*$ would be stronger, and so the galaxy will self-regulate at the same $\langle \dot{M}_* \rangle$. Thus, the instantaneous self-regulation (the Kennicutt-Schmidt relation) is not the same as self-regulation of cosmological growth (e.g. the stellar mass-halo mass relation).

10 SUMMARY & CONCLUSIONS

10.1 Overview

We present the FIRE-2 simulations (Table 1), a suite of cosmological simulations of galaxy formation using the FIRE physics modules in the GIZMO code. The FIRE-2 suite represents an update over FIRE-1, primarily in the use of a newer, more accurate hydrodynamic method, together with other numerical improvements to the physics algorithms and resolution – but no significant change in the actual *physics* simulated (§ 2). To our knowledge, this includes the highest-resolution simulations (to date) run to $z = 0$ with fully baryonic physics, at both the dwarf and MW mass scales (reaching sub-pc spatial and $\sim 100\,M_\odot$ mass scales; see Table 3 for details).

In all properties investigated here, our primary conclusions from FIRE-1 appear robust to these improvements in numerical accuracy (see § 3, Figs. 5-7), with typically less than factor of a few changes in global galaxy properties such as stellar mass, metallicity, and star formation rates. With *explicit* treatment of stellar feedback from SNe Types Ia & II, OB & AGB stellar winds, radiation pressure (UV, optical, and IR), photo-ionization and photo-electric heating, and explicit resolution of the multi-phase ISM, galactic winds emerge naturally, producing galaxies with morphologies (Figs. 1 & 3), internal ISM structure (Fig. 2), flat or rising star formation histories and flat rotation curves (Figs. 4-5), metallicities (Fig. 4), and stellar masses (Fig. 7) in good agreement with observations. This is remarkable, considering that there is no fine-tuning or direct calibration of any parameters in the simulations to match these observations. Each of these properties has been investigated in more detail in previous work, described in § 1.

It is worth noting that, as previous FIRE papers have pointed out (Hopkins et al. 2014; Ma et al. 2016d; Wetzel et al. 2016), there is no tension between even very thin stellar disks at $z = 0$ (e.g. Fig. 1), and strong stellar feedback which drives bursty star formation and strong galactic outflows at high redshift, strongly suppressing the stellar masses at $z = 0$ relative to no-feedback models. This owes to a combination of resolved venting of hot winds through a multi-phase ISM, which remove mass without disturbing galaxy morphologies, and a rapid transition from the bursty star-forming mode and more "gentle" steady-state thin-disk mode as gas fractions and specific star formation rates decline in massive, low-redshift halos (see Hayward & Hopkins 2017; Faucher-Giguère 2017; Ma et al. 2016d).

We have provided a complete algorithmic description of all aspects of the FIRE simulations, sufficient to reproduce our results here (see § 2 & A-F). We also make our hydrodynamics+gravity code GIZMO, and all initial conditions for the runs here, public.

We have considered an extensive study varying numerical and physical aspects of our simulations, to identify the most important "ingredients" and physical effects in the simulations. Table 2 summarizes our conclusions; below we note the most important.

10.2 Resolution

(i) Mass resolution is most important for convergence in our Lagrangian methods; in § 4.1 we present a number of new resolution criteria relevant for simulations of resolved galaxy formation, phase structure, and galactic winds. For example, for dwarf galaxies stellar masses are well-converged with just a few star particles per galaxy (Fig. 8). Metallicity and the shape of the SFH converge more slowly, requiring ~ 100 star particles to sample the self-enrichment history and a baryonic particle mass $\lesssim 10^{-6} M_{\rm halo}$ (Eq. 7) to avoid numerically enhanced burstiness (from single star particles representing "too many" SNe at once). Resolving phase structure in the ISM and gravitational fragmentation requires resolving the Toomre mass in the ISM, i.e. the largest GMCs, which dominate the galactic SFR (Eq. 5; Fig. 13). Morphology converges rapidly for dwarfs; for massive galaxies, it appears to depend on galaxy formation history – some massive systems are disky at all resolution levels, others (where the disk forms somewhat later) are more sensitive (Fig. 9). Details of galactic winds, in particular hot gas "venting" and recycling in the CGM, converge most slowly (of properties we consider), with true convergence requiring the ability to resolve the cooling radii of individual SNe (Eq. 6).

(ii) Force softenings for *collisionless* (DM+stellar) particles have little effect on our predictions (including galaxy baryonic properties, DM halo mass profiles, sub-structure mass and velocity distribution functions), provided they are chosen within a broad range (§ 4.2, Figs. 15-19). For DM this is between $\sim (0.005 - 0.5)\, r_{0.06}$, where $r_{0.06}$ is approximately the radius containing ~ 200 DM particles at the halo center at $z = 0$ (§ 4.2.3-4.2.5).

(iii) We show that the radius $r_{0.06}$ is the radius to which DM profiles are converged to $\sim 10\%$ or better – this is significantly smaller than the commonly quoted convergence radius of Power et al. (2003), owing to more accurate integration and timestepping, and smoother kernels (Figs. 11-12). At these spatial and mass resolution scales, N-body heating is negligible (Eq. 10-11). Owing to this insensitivity, it makes little difference whether we adopt constant softenings or fully-adaptive softenings (set to a fixed multiple of the local inter-particle separation) for collisionless particles.

(iv) Force softenings for *gas* should be set adaptively, so that the *same* mass distribution is being treated in the hydrodynamic equations and gravity equations (§ 4.2). With such a choice, turbulent fragmentation can be resolved down to small spatial scales, and no additional resolution criteria need be applied; of course, the meaning of "spatial resolution" is adaptive, so Table 3 summarizes the effective spatial resolution of our simulations in different regimes. In dense, star-forming gas, the effective resolution reaches $\sim 1 - 10\,$pc in MW-mass systems and $\sim 0.1 - 1\,$pc in dwarfs. Because the medium is super-sonically turbulent, convergence in turbulent fragmentation and galactic star formation does *not* require that the *thermal* Jeans mass/length be resolved.

(v) If, for some reason, constant gravitational force softening for gas is desired, we show (Figs. 15-16) that it should be chosen sufficiently small that (1) the *Toomre* length scale (molecular gas disk scale height) is resolved ($\epsilon \ll 100\,$pc), and (2) the minimum "threshold" density for star formation is resolved *gravitationally*, so that actual self-gravitating regions are what form stars ($\epsilon \lesssim 7\,{\rm pc}\, m_{i,1000}^{1/3}\, (n_{\rm SF,\,min}/100\,{\rm cm^{-3}})^{-1/3}$). The latter criterion is often violated in the literature – but this means that in such simulations star formation is driven (incorrectly) entirely by *global* (galaxy-wide) contraction or shock-compression of gas, and has nothing to do with local fragmentation and collapse of GMCs.

(vi) In addition to standard timestep criteria, guaranteeing proper time-resolution of feedback requires a limiter (Eq. 12) such that star particles cannot "skip" stages of stellar evolution or experience $\gg 1$ SNe per particle per timestep (§ 4.3).

(vii) Moreover, if one adopts adaptive gravitational softening for collisionless particles, we show that standard timesteps based on the acceleration are numerically unstable. The reason is that the accelerations and potential now depend explicitly on the local particle configuration which changes as particles move through one another, not just on the total potential gradient; this means a Courant-like condition *must* be used to ensure stability (Fig. 20; Eq. 13).

10.3 Hydrodynamic Methods

(i) Comparing our new mesh-free finite-volume Godunov hydrodynamic methods ("MFM") to SPH (§ 5), we show that at all resolution levels, the properties of dwarf and high-redshift galaxies are remarkably insensitive to the details of the hydrodynamic solver (Fig. 21). However, for MW-mass and larger galaxies at late times, when a "hot halo" develops, the details of the hydrodynamics become important for the late-time steady-state SFR of the galaxies, and their gas disk sizes (and subsequent circular velocity profiles). Similar conclusions have been reached in previous comparisons of moving-mesh and fixed-grid codes to SPH, but with much simpler treatments of feedback (Torrey et al. 2012; Sijacki et al. 2012; Kereš et al. 2012; Vogelsberger et al. 2012). We show that the gross morphology and instantaneous SFR given some ISM gas supply are not particularly sensitive (Figs. 23-24). Instead, the difference owes to the mixing and recycling of galactic winds in the CGM and initial accretion and shocking of hot gas, and its subsequent non-linear effect on cooling and re-accretion (Figs. 25-26). SPH – even our state-of-the-art "modern" SPH with a pressure-energy formulation, "inviscid" artificial viscosity prescription, artificial conductivity terms, higher-order kernels, and more improvements – produces spurious "blobs" in the CGM which do not mix, and "smears out" accretion shocks and suppresses sub-sonic turbulence, leading to more cold ($\sim 10 - 1000\,$K) gas and less hot, dense gas in the CGM. We explicitly show that by adding stronger "artificial conductivity" to our SPH runs, we can essentially reproduce our MFM results (Fig. 22). This clearly indicates that details of numerical diffusion are important in the CGM of "hot halos."

(ii) We argue that it is never appropriate to use "artificial pressure floors" when a sink-particle model like ours for star formation (which identifies self-gravitating regions to turn into stars) is already present (§ 6). However, for modest values sometimes used in the literature, we show such a floor has weak effects – but making the pressure floors too large produces clear numerical artifacts and prevents proper physical convergence in treating turbulent fragmentation (Fig. 27).

(iii) In finite-mass numerical methods without explicit advective fluxes (MFM), passive scalars (e.g. metals) remain locked (by default) to gas elements where they are first injected; in methods with such fluxes (e.g. AMR) there is a significant numerical mixing between elements. At finite resolution, these tend to under and over-estimate (respectively) true physical mixing by un-resolved turbu-

lent eddies and microphysical diffusion between neighbor cells. For AMR codes reducing the error requires higher resolution; in MFM one can add an explicit metal diffusivity. However, while this can be important for predictions of detailed abundance pattern distributions within galaxies, we show that this has essentially no effect on any properties studied here (§ 7.2; Fig 30-31), regardless of the numerical implementation (Fig. 32).

(iv) Given the sensitivity of the CGM in hot halos to fluid mixing details, it is likely that additional physics – magnetic fields, anisotropic Spitzer-Braginskii conduction and viscosity, cosmic ray transport – may produce larger changes in the late-time cooling rates from hot halos than the difference between hydro solvers.

10.4 Cooling, Star Formation, & Feedback

(i) Consistent with previous work (Piontek & Steinmetz 2009; Hopkins et al. 2011, 2012c; Glover & Clark 2012; Hu et al. 2016), we find that the *details* of radiative cooling and gas chemistry generally have little effect on galaxy dynamics, star formation, and galactic winds (§ 7; Fig. 27-28). This is because almost all gas within the ISM has cooling times much shorter than dynamical times, so the exact cooling time and/or temperature is dynamically irrelevant. This is especially true in the "cold" and "cool" phases of the ISM/CGM/IGM ($T \lesssim 10^5$ K). Detailed variations in yields, which species are tracked, and numerical metal-mixing produce correspondingly weak effects.

(ii) However, in hot gas ($T \gg 10^6$ K) in the CGM or SNe-heated bubbles, the cooling time can be longer than the dynamical time, and so high-temperature metal-line cooling has a non-negligible effect on the phase structure of the CGM (Fig. 29) and cooling onto (hence SF in) the galaxy from the halo (Fig. 28; see also Choi & Nagamine 2009; Schaye et al. 2010). To leading order the total metallicity is the important quantity here, while detailed abundance-ratio variations have a much smaller effect.

(iii) As we have repeatedly shown in previous work, provided the largest (Toomre) scales of fragmentation are resolved (*gravitationally* and *hydrodynamically*), and star formation occurs in that fragmenting gas, the resolution-scale star formation prescription has essentially no detectable effects on our predictions for galaxy-integrated SFRs, stellar masses, sizes, positions on the Schmidt-Kennicutt law (both galaxy-wide and spatiall-resolved), galactic winds, and more (see § 8, Figs. 33-34, and Saitoh et al. 2008; Shetty & Ostriker 2008; Hopkins et al. 2011, 2013d; Kim et al. 2011; Agertz et al. 2013). This includes orders-of-magnitude variations in the density threshold and resolution-scale "rate per free-fall time." The key criteria are that these densities can (of course) be resolved, and that the SF occur only in self-gravitating gas above the mean galaxy density, so that the SF is automatically clustered (since the natural clustering of star formation non-linearly influences the likelihood of e.g. SNe bubbles overlapping and driving superwinds). Of course, predictions *within the dense gas* (corresponding to the densities where our sub-grid SF models take over) will be sensitive to the exact choices made; our default model is chosen based on previous work arguing it best reproduced observations of dense gas tracers (HCN and CO(3-2)) in idealized (non-cosmological) simulations (Hopkins et al. 2013e).

(iv) Provided the above criteria are met, SF is instead *feedback-regulated*. Uniformly increasing/decreasing the strength of feedback directly shifts both the integrated stellar mass, and position of galaxies on the KS law (Fig. 35).

(v) Stellar mass-loss (OB/AGB winds) is primarily important as a late-time fuel source for star formation (§ 9; Figs. 36-37); its net effect is to *increase*, not decrease, late-time SFRs, especially in massive galaxies.

(vi) Radiative feedback has a strong effect regulating the *instantaneous* SFRs of galaxies (i.e. the Kennicutt-Schmidt law), though less so on cosmologically-averaged SFRs (which are also regulated by the availability of fresh gas from IGM accretion).

(vii) The most important form of radiative feedback for dwarfs is *external* radiative feedback, i.e. the UV background. In Paper III, we show that removing the UVB has a much larger effect on the stellar masses of dwarfs than removing internal radiative feedback, even at mass scales as large as $M_* \gtrsim 10^7\,M_\odot$. We will explore these effects further in future work.

(viii) On cosmological scales, SNe (specifically core-collapse and prompt Ia) are the most important feedback mechanism regulating galaxy properties. Without SNe (even given other feedback mechanisms), galaxies (especially dwarfs) form too many stars, and form these stars far too early. In Paper II, we therefore investigate numerical SNe treatments in more detail.

10.5 Summary of Ingredients

To summarize, we find that the following criteria are essential for converged, physically realistic high-resolution zoom-in simulations of galaxy formation:

(i) The Toomre mass/length is mass and force-resolved; disk scale heights and at least the largest scales of fragmentation are resolved, with gravitational softening able to follow the mass.

(ii) In addition to standard (optically thin, primordial) cooling, some accounting for self-shielding and low-temperature ($T \ll 10^4$ K) cooling is included (to enable fragmentation), and high-temperature metal-line cooling is included to account for faster cooling in metal-enriched halo gas.

(iii) Star formation is restricted to gas at (force-resolved) densities significantly larger than the mean (ideally, to self-gravitating gas with a sink-particle type approach), so that it will occur in said fragments and therefore naturally be clustered, as observed.

(iv) Stellar mass loss is included, to provide a continuous additional fuel supply to the galaxy.

(v) Radiative feedback is included, particularly heating from the meta-galactic UV background, as well as photo-ionization heating and single-scattering photon momentum (radiation pressure) from stars in the simulation.

(vi) SNe (Ia & II) are included, with a coupling algorithm that carefully ensures manifest conservation of energy, mass, and momentum. More importantly, a careful accounting of both the energy and momentum budget of the coupled terms that properly treats whether the coupled radii are inside or outside the cooling radius, is necessary to obtain converged solutions (see Paper II).

(vii) All feedback quantities follow standard stellar evolution models for a standard IMF.

10.6 Future Work

In companion papers, we will study in more detail how mechanical feedback (SNe & stellar mass-loss; Paper II) and radiative feedback (photo-heating & photon momentum; Paper III) influence our predictions – both the details of their physics *and* numerical implementations. Because we argue that feedback is more important than most numerical details, it is extremely important to treat it as accurately as possible.

As discussed above, more work is clearly warranted to investigate how additional fluid microphysics (e.g. magnetic fields) alters

fluid mixing and subsequent cooling from the CGM. A major motivation of our switch to the new hydrodynamics solver in FIRE-2 is that it allows us to incorporate such physics in future work. A preliminary investigation of some of these physics is presented in Su et al. (2016). However this was primarily focused on the ISM inside galaxies; detailed, higher-resolution CGM studies are clearly necessary.

In this paper, we study only systems with halo masses $\lesssim 2\times10^{12}\,M_{\odot}$. This is because it is widely believed that feedback from supermassive black holes (not included in our simulations here) is critical to explain the properties (especially the quenching of star formation and further quiescence) of these "red and dead" systems (see Hopkins et al. 2006; Croton et al. 2006). Our preliminary studies from FIRE-1 support the idea that stellar feedback alone cannot explain all the observed properties of these most massive galaxies (Hopkins et al. 2014; Feldmann et al. 2016b). In future work, we study the effects of black hole feedback on galaxy properties.

Finally, we focus here almost exclusively on numerical studies. A series of papers will present the scientific predictions of these simulations for current and future observations.

ACKNOWLEDGMENTS

Support for PFH and co-authors was provided by an Alfred P. Sloan Research Fellowship, NASA ATP Grant NNX14AH35G, and NSF Collaborative Research Grant #1411920 and CAREER grant #1455342. ARW was supported by a Caltech-Carnegie Fellowship, in part through the Moore Center for Theoretical Cosmology and Physics at Caltech. CAFG was supported by NSF through grants AST-1412836 and AST-1517491, and by NASA through grant NNX15AB22G. The Flatiron Institute is supported by the Simons Foundation. MBK and AF were partially supported by the NSF through grant AST-1517226. MBK also acknowledges support from NASA through HST theory grants (programs AR-12836, AR-13888, AR-13896, and AR-14282) awarded by STScI. DK was supported by NSF Grant AST1412153 and a Cottrell Scholar Award from the Research Corporation for Science Advancement. Numerical calculations were run on the Caltech compute cluster "Wheeler," allocations TG-AST130039 & TG-AST150080 granted by the Extreme Science and Engineering Discovery Environment (XSEDE) supported by the NSF, and the NASA HEC Program through the NAS Division at Ames Research Center and the NCCS at Goddard Space Flight Center.

REFERENCES

Abel, T., Hahn, O., & Kaehler, R. 2012, MNRAS, 427, 61
Achterberg, A. 1981, A&A, 98, 161
Agertz, O., & Kravtsov, A. V. 2016, ApJ, 824, 79
Agertz, O., Kravtsov, A. V., Leitner, S. N., & Gnedin, N. Y. 2013, ApJ, 770, 25
Agertz, O., et al. 2007, MNRAS, 380, 963
Aguirre, A., Hernquist, L., Schaye, J., Weinberg, D. H., Katz, N., & Gardner, J. 2001, ApJ, 560, 599
Anglés-Alcázar, D., Davé, R., Faucher-Giguère, C.-A., Özel, F., & Hopkins, P. F. 2017, MNRAS, 464, 2840
Anglés-Alcázar, D., Davé, R., Özel, F., & Oppenheimer, B. D. 2014, ApJ, 782, 84
Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., Hopkins, P. F., Quataert, E., & Murray, N. 2016, MNRAS, in press, arXiv:1610.08523
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Athanassoula, E., Fady, E., Lambert, J. C., & Bosma, A. 2000, MNRAS, 314, 475
Bagla, J. S., & Khandai, N. 2009, MNRAS, 396, 2211
Barnes, J. E. 2012, MNRAS, 425, 1104
Bate, M. R. 2012, MNRAS, 419, 3115
Bate, M. R., & Burkert, A. 1997, MNRAS, 288, 1060
Bauer, A., & Springel, V. 2012, MNRAS, 423, 3102
Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2012, ApJ, in press, arXiv:1207.6105
Benincasa, S. M., Wadsley, J., Couchman, H. M. P., & Keller, B. W. 2016, MNRAS, in press, arXiv:1607.05795
Berezinskii, V. S., Bulanov, S. V., Dogiel, V. A., & Ptuskin, V. S. 1990, Astrophysics of cosmic rays (Amsterdam: North-Holland, 1990, edited by Ginzburg, V.L.)
Blitz, L., & Rosolowsky, E. 2005, in Astrophysics and Space Science Library, Vol. 327, The Initial Mass Function 50 Years Later; Springer, Dordrecht, ed. E. Corbelli, F. Palla, & H. Zinnecker, 287
Bolatto, A. D., Leroy, A. K., Rosolowsky, E., Walter, F., & Blitz, L. 2008, ApJ, 686, 948
Bouché, N., et al. 2010, ApJ, 718, 1001
Bournaud, F., Elmegreen, B. G., Teyssier, R., Block, D. L., & Puerari, I. 2010, MNRAS, 409, 1088
Bournaud, F., et al. 2011, ApJ, 730, 4
Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, eprint arXiv, 0903, 3041
Braginskii, S. I. 1965, Reviews of Plasma Physics, 1, 205
Brook, C. B., Di Cintio, A., Knebe, A., Gottlöber, S., Hoffman, Y., Yepes, G., & Garrison-Kimmel, S. 2014, ApJL, 784, L14
Brook, C. B., et al. 2011, MNRAS, 415, 1051
Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80
Bryan, G. L., et al. 2014, ApJS, 211, 19
Byerly, Z. D., Adelstein-Lelbach, B., Tohline, J. E., & Marcello, D. C. 2014, ApJS, 212, 23
Chan, T. K., Kereš, D., Oñorbe, J., Hopkins, P. F., Muratov, A. L., Faucher-Giguère, C.-A., & Quataert, E. 2015, MNRAS, 454, 2981
Chen, Y.-M., Tremonti, C. A., Heckman, T. M., Kauffmann, G., Weiner, B. J., Brinchmann, J., & Wang, J. 2010, AJ, 140, 445
Chevalier, R. A., & Gardner, J. 1974, ApJ, 192, 457
Chiaki, G., & Yoshida, N. 2015, MNRAS, 451, 3955
Chieffi, A., & Limongi, M. 2004, ApJ, 608, 405
Choi, E., Ostriker, J. P., Naab, T., Oser, L., & Moster, B. P. 2015, MNRAS, 449, 4105
Choi, J., & Nagamine, K. 2009, MNRAS, 393, 1595
Christensen, C. R., Davé, R., Governato, F., Pontzen, A., Brooks, A., Munshi, F., Quinn, T., & Wadsley, J. 2016, ApJ, 824, 57
Cioffi, D. F., McKee, C. F., & Bertschinger, E. 1988, ApJ, 334, 252
Ciotti, L., & Ostriker, J. P. 2001, ApJ, 551, 131
Ciotti, L., Ostriker, J. P., & Proga, D. 2009, ApJ, 699, 89
Coil, A. L., Weiner, B. J., Holz, D. E., Cooper, M. C., Yan, R., & Aird, J. 2011, ApJ, 743, 46
Colbrook, M. J., Ma, X., Hopkins, P. F., & Squire, J. 2016, MNRAS, in press, arXiv:1610.06590
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, ApJ, 647, 201
Crain, R. A., et al. 2015, MNRAS, 450, 1937

Croton, D. J., et al. 2006, MNRAS, 365, 11
Cullen, L., & Dehnen, W. 2010, MNRAS, 408, 669
Davé, R., Finlator, K., & Oppenheimer, B. D. 2006, MNRAS, 370, 273
—. 2011a, MNRAS, 416, 1354
—. 2012, MNRAS, 421, 98
Davé, R., Oppenheimer, B. D., & Finlator, K. 2011b, MNRAS, 415, 11
Davé, R., Thompson, R., & Hopkins, P. F. 2016, MNRAS, 462, 3265
Dehnen, W. 2001, MNRAS, 324, 273
Dehnen, W., & Aly, H. 2012, MNRAS, 425, 1068
Dobbs, C. L., Burkert, A., & Pringle, J. E. 2011, MNRAS, 413, 528
Draine, B. T., & Woods, D. T. 1991, ApJ, 383, 621
Durier, F., & Dalla Vecchia, C. 2012, MNRAS, 419, 465
Eisenstein, D. J., & Hut, P. 1998, ApJ, 498, 137
El-Badry, K., Wetzel, A., Geha, M., Hopkins, P. F., Kereš, D., Chan, T. K., & Faucher-Giguère, C.-A. 2016, ApJ, 820, 131
Elmegreen, B. G. 2007, ApJ, 668, 1064
Enßlin, T., Pfrommer, C., Miniati, F., & Subramanian, K. 2011, A&A, 527, A99
Evans, N. J., et al. 2009, ApJS, 181, 321
Evans, II, N. J. 1999, ARA&A, 37, 311
Evrard, A. E. 1988, MNRAS, 235, 911
Faucher-Giguère, C.-A. 2017, ArXiv e-prints
Faucher-Giguère, C.-A., Feldmann, R., Quataert, E., Kereš, D., Hopkins, P. F., & Murray, N. 2016, MNRAS, 461, L32
Faucher-Giguère, C.-A., Hopkins, P. F., Keres, D., Muratov, A. L., Quataert, E., & Murray, N. 2015, MNRAS, 449, 987
Faucher-Giguère, C.-A., Kereš, D., Dijkstra, M., Hernquist, L., & Zaldarriaga, M. 2010, ApJ, 725, 633
Faucher-Giguère, C.-A., Kereš, D., & Ma, C.-P. 2011, MNRAS, 417, 2982
Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., & Hernquist, L. 2009, ApJ, 703, 1416
Faucher-Giguère, C.-A., & Quataert, E. 2012, MNRAS, 425, 605
Faucher-Giguère, C.-A., Quataert, E., & Hopkins, P. F. 2013, MNRAS, 433, 1970
Federrath, C. 2015, MNRAS, 450, 4035
Federrath, C., Banerjee, R., Clark, P. C., & Klessen, R. S. 2010, ApJ, 713, 269
Federrath, C., & Klessen, R. S. 2012, ApJ, 761, 156
—. 2013, ApJ, 763, 51
Federrath, C., Klessen, R. S., & Schmidt, W. 2008, ApJL, 688, L79
Feldmann, R. 2013, MNRAS, 433, 1910
—. 2015, MNRAS, 449, 3274
Feldmann, R., & Gnedin, N. Y. 2011, ApJL, 727, L12+
Feldmann, R., Hopkins, P. F., Quataert, E., Faucher-Giguère, C.-A., & Kereš, D. 2016a, MNRAS, 458, L14
Feldmann, R., Quataert, E., Hopkins, P. F., Faucher-Giguère, C.-A., & Kereš, D. 2016b, MNRAS, in press, arXiv:1610.02411
Felice, G. M., & Kulsrud, R. M. 2001, ApJ, 553, 198
Ferland, G. J., Korista, K. T., Verner, D. A., Ferguson, J. W., Kingdon, J. B., & Verner, E. M. 1998, PASP, 110, 761
Few, C. G., Dobbs, C., Pettitt, A., & Konstandin, L. 2016, MNRAS, 460, 4382
Fitts, A., et al. 2016, MNRAS, in press, arXiv:1611.02281
Frenk, C. S., et al. 1999, ApJ, 525, 554
Gaburov, E., & Nitadori, K. 2011, MNRAS, 414, 129
Gaggero, D., Urbano, A., Valli, M., & Ullio, P. 2015, PhRvD, 91, 083012
Gan, Z., Yuan, F., Ostriker, J. P., Ciotti, L., & Novak, G. S. 2014, ApJ, 789, 150
García-Senz, D., Cabezón, R. M., & Escartín, J. A. 2012, A&A, 538, A9
Garrison-Kimmel, S., Bullock, J. S., Boylan-Kolchin, M., & Bardwell, E. 2016, MNRAS, in press, arXiv:1603.04855
Glover, S. C. O., & Clark, P. C. 2012, MNRAS, 421, 9
Gnedin, N. Y., & Abel, T. 2001, New Astronomy, 6, 437
Goldsmith, P. F., & Langer, W. D. 1978, ApJ, 222, 881
Gong, H., & Ostriker, E. C. 2013, ApJS, 204, 8
Governato, F., et al. 2010, Nature, 463, 203
Grudić, M. Y., Hopkins, P. F., Faucher-Giguère, C.-A., Quataert, E., Murray, N., & Kereš, D. 2016, MNRAS, in press, arXiv:1612.05635
Guedes, J., Callegari, S., Madau, P., & Mayer, L. 2011, ApJ, 742, 76
Guo, F., & Oh, S. P. 2008, MNRAS, 384, 251
Guo, Q., White, S., Li, C., & Boylan-Kolchin, M. 2010, MNRAS, 404, 1111
Guszejnov, D., & Hopkins, P. F. 2015, MNRAS, 450, 4137
—. 2016, MNRAS, 459, 9
Guszejnov, D., Krumholz, M. R., & Hopkins, P. F. 2016, MNRAS, 458, 673
Habing, H. J. 1968, Bulletin of the Astronomical Institutes of the Netherlands, 19, 421
Hafen, Z., et al. 2016, MNRAS, in press, arXiv:1608.05712
Hahn, O., & Abel, T. 2011, MNRAS, 415, 2101
Hahn, O., Teyssier, R., & Carollo, C. M. 2010, MNRAS, 405, 274
Haid, S., Walch, S., Naab, T., Seifried, D., Mackey, J., & Gatto, A. 2016, MNRAS, in press, arXiv:1604.04395
Harper-Clark, E., & Murray, N. 2011, in Computational Star Formation; Cambridge University Press, ed. J. Alves, B. G. Elmegreen, J. M. Girart & V. Trimble, Vol. 270 (Cambridge, UK: Cambridge University Press), 235–238
Hayward, C. C., & Hopkins, P. F. 2017, MNRAS, 465, 1682
Hayward, C. C., Torrey, P., Springel, V., Hernquist, L., & Vogelsberger, M. 2014, MNRAS, 442, 1992
Heckman, T. M., Lehnert, M. D., Strickland, D. K., & Armus, L. 2000, ApJS, 129, 493
Heitmann, K., et al. 2008, Computational Science and Discovery, 1, 015003
Hennebelle, P., & Chabrier, G. 2008, ApJ, 684, 395
Holman, G. D., Ionson, J. A., & Scott, J. S. 1979, ApJ, 228, 576
Hopkins, P. F. 2012a, MNRAS, 423, 2016
—. 2012b, MNRAS, 423, 2037
—. 2013a, MNRAS, 428, 2840
—. 2013b, MNRAS, 430, 1653
—. 2013c, MNRAS, 433, 170
—. 2013d, MNRAS, 428, 1950
—. 2015, MNRAS, 450, 53
—. 2016a, MNRAS, 462, 576
—. 2016b, MNRAS, in press, arXiv:1602.07703
Hopkins, P. F., & Christiansen, J. L. 2013, ApJ, 776, 48
Hopkins, P. F., Cox, T. J., Hernquist, L., Narayanan, D., Hayward, C. C., & Murray, N. 2013a, MNRAS, 430, 1901
Hopkins, P. F., Hernquist, L., Cox, T. J., Di Matteo, T., Robertson, B., & Springel, V. 2006, ApJS, 163, 1
Hopkins, P. F., Hernquist, L., Martini, P., Cox, T. J., Robertson, B., Di Matteo, T., & Springel, V. 2005, ApJL, 625, L71

Hopkins, P. F., Keres, D., Murray, N., Quataert, E., & Hernquist, L. 2012a, MNRAS, 427, 968
Hopkins, P. F., Keres, D., Onorbe, J., Faucher-Giguère, C.-A., Quataert, E., Murray, N., & Bullock, J. S. 2014, MNRAS, 445, 581
Hopkins, P. F., Kereš, D., & Murray, N. 2013b, MNRAS, 432, 2639
Hopkins, P. F., Kereš, D., Murray, N., Hernquist, L., Narayanan, D., & Hayward, C. C. 2013c, MNRAS, 433, 78
Hopkins, P. F., & Lee, H. 2016, MNRAS, 456, 4174
Hopkins, P. F., Narayanan, D., & Murray, N. 2013d, MNRAS, 432, 2647
Hopkins, P. F., Narayanan, D., Murray, N., & Quataert, E. 2013e, MNRAS, 433, 69
Hopkins, P. F., Quataert, E., & Murray, N. 2011, MNRAS, 417, 950
—. 2012b, MNRAS, 421, 3522
—. 2012c, MNRAS, 421, 3488
Hopkins, P. F., & Raives, M. J. 2016, MNRAS, 455, 51
Hopkins, P. F., Torrey, P., Faucher-Giguère, C.-A., Quataert, E., & Murray, N. 2016, MNRAS, 458, 816
Hu, C.-Y., Naab, T., Walch, S., Glover, S. C. O., & Clark, P. C. 2016, MNRAS, 458, 3528
Hu, C.-Y., Naab, T., Walch, S., Moster, B. P., & Oser, L. 2014, MNRAS(arXiv:1402.1788)
Hubber, D. A., Allison, R. J., Smith, R., & Goodwin, S. P. 2013, MNRAS, 430, 1599
Hummels, C. B., Bryan, G. L., Smith, B. D., & Turk, M. J. 2013, MNRAS, 430, 1548
Iannuzzi, F., & Athanassoula, E. 2013, MNRAS, 436, 1161
Iannuzzi, F., & Dolag, K. 2011, MNRAS, 417, 2846
Iffrig, O., & Hennebelle, P. 2015, A&A, 576, A95
Iwamoto, K., Brachwitz, F., Nomoto, K., Kishimoto, N., Umeda, H., Hix, W. R., & Thielemann, F.-K. 1999, ApJS, 125, 439
Izzard, R. G., Tout, C. A., Karakas, A. I., & Pols, O. R. 2004, MNRAS, 350, 407
Jungwiert, B., Combes, F., & Palouš, J. 2001, Astronomy and Astrophysics, 376, 85
Kannan, R., Stinson, G. S., Macciò, A. V., Brook, C., Weinmann, S. M., Wadsley, J., & Couchman, H. M. P. 2013, MNRAS, in press, arxiv:1302.2618
Katz, N. 1992, ApJ, 391, 502
Katz, N., Weinberg, D. H., & Hernquist, L. 1996, ApJS, 105, 19
Katz, N., & White, S. D. M. 1993, ApJ, 412, 455
Kennicutt, Jr., R. C. 1998, ApJ, 498, 541
Kereš, D., Katz, N., Davé, R., Fardal, M., & Weinberg, D. H. 2009, MNRAS, 396, 2332
Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
Kereš, D., Vogelsberger, M., Sijacki, D., Springel, V., & Hernquist, L. 2012, MNRAS, 425, 2027
Kim, C.-G., Kim, W.-T., & Ostriker, E. C. 2011, ApJ, 743, 25
Kim, C.-G., & Ostriker, E. C. 2015, ApJ, 802, 99
Kim, C.-G., Ostriker, E. C., & Kim, W.-T. 2013a, ApJ, 776, 1
Kim, J.-h., et al. 2013b, ApJ, in press, arXiv:1308.2669
Kimm, T., & Cen, R. 2014, ApJ, 788, 121
Kitsionas, S., et al. 2009, A&A, 508, 541
Klypin, A., Gottlöber, S., Kravtsov, A. V., & Khokhlov, A. M. 1999, ApJ, 516, 530
Kratter, K. M., & Lodato, G. 2016, ARA&A, in press, arXiv:1603.01280
Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., Klypin, A. A., Gottlöber, S., Allgood, B., & Primack, J. R. 2004, ApJ, 609, 35
Kravtsov, A. V., Klypin, A. A., & Khokhlov, A. M. 1997, ApJS, 111, 73
Kroupa, P. 2001, MNRAS, 322, 231
Krumholz, M. R., & Gnedin, N. Y. 2011, ApJ, 729, 36
Krumholz, M. R., Klein, R. I., & McKee, C. F. 2007, ApJ, 656, 959
—. 2011, ApJ, 740, 74
Kulsrud, R., & Pearce, W. P. 1969, ApJ, 156, 445
Lanson, N., & Vila, J.-P. 2008, SIAM J. Numer. Anal., 46, 1912
Lee, A. T., Cunningham, A. J., McKee, C. F., & Klein, R. I. 2014, ApJ, 783, 50
Leitherer, C., et al. 1999, ApJS, 123, 3
Leitner, S. N., & Kravtsov, A. V. 2011, ApJ, 734, 48
Levermore, C. D. 1984, Journal of Quantitative Spectroscopy and Radiative Transfer, 31, 149
Li, M., Ostriker, J. P., Cen, R., Bryan, G. L., & Naab, T. 2015, ApJ, 814, 4
Li, Y., Mac Low, M.-M., & Klessen, R. S. 2005, ApJL, 620, L19
Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119
Lopez, L. A., Krumholz, M. R., Bolatto, A. D., Prochaska, J. X., & Ramirez-Ruiz, E. 2011, ApJ, 731, 91
Lukat, G., & Banerjee, R. 2016, New Astronomy, 45, 14
Luo, H., Baum, J. D., & Löhner, R. 2008, Journal of Computational Physics, 227, 8875
Ma, X., Hopkins, P. F., Faucher-Giguère, C.-A., Zolman, N., Muratov, A. L., Kereš, D., & Quataert, E. 2016a, MNRAS, 456, 2140
Ma, X., Hopkins, P. F., Feldmann, R., Torrey, P., Faucher-Giguère, C.-A., & Keres, D. 2016b, MNRAS, in press, arXiv:1610.03498
Ma, X., Hopkins, P. F., Kasen, D., Quataert, E., Faucher-Giguère, C.-A., Kereš, D., Murray, N., & Strom, A. 2016c, MNRAS, 459, 3614
Ma, X., Hopkins, P. F., Wetzel, A. R., Kirby, E. N., Angles-Alcazar, D., Faucher-Giguère, C.-A., Keres, D., & Quataert, E. 2016d, MNRAS, in press, arXiv:1608.04133
Ma, X., Kasen, D., Hopkins, P. F., Faucher-Giguère, C.-A., Quataert, E., Kereš, D., & Murray, N. 2015, MNRAS, 453, 960
Macciò, A. V., Stinson, G., Brook, C. B., Wadsley, J., Couchman, H. M. P., Shen, S., Gibson, B. K., & Quinn, T. 2012, ApJL, 744, L9
Mannucci, F., Della Valle, M., & Panagia, N. 2006, MNRAS, 370, 773
Manuel, Z.-A., Enrique, V.-S., Bastian, K., Robi, B., & Lee, H. 2016, MNRAS, in press, arXiv:1606.05343
Marigo, P. 2001, A&A, 370, 194
Maron, J. L., & Howes, G. G. 2003, ApJ, 595, 564
Martin, C. L. 1999, ApJ, 513, 156
—. 2006, ApJ, 647, 222
Martin, C. L., Scannapieco, E., Ellison, S. L., Hennawi, J. F., Djorgovski, S. G., & Fournier, A. P. 2010, ApJ, 721, 174
Martizzi, D., Faucher-Giguère, C.-A., & Quataert, E. 2015, MNRAS, 450, 504
Meijerink, R., & Spaans, M. 2005, A&A, 436, 397
Merritt, D. 1996, AJ, 111, 2462
Mitra, S., Davé, R., Simha, V., & Finlator, K. 2017, MNRAS, 464, 2766
Mocz, P., Vogelsberger, M., Sijacki, D., Pakmor, R., & Hernquist, L. 2014, MNRAS, 437, 397
Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121

Moster, B. P., Somerville, R. S., Maulbetsch, C., van den Bosch, F. C., Macciò, A. V., Naab, T., & Oser, L. 2010, ApJ, 710, 903
Müller, E., & Steinmetz, M. 1995, Computer Physics Communications, 89, 45
Muratov, A. L., Gnedin, O. Y., Gnedin, N. Y., & Zemp, M. 2013, ApJ, 772, 106
Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., Hopkins, P. F., Quataert, E., & Murray, N. 2015, MNRAS, 454, 2691
Muratov, A. L., et al. 2016, MNRAS, in press, arXiv:1606.09252
Murray, N. 2011, ApJ, 729, 133
Murray, N., Quataert, E., & Thompson, T. A. 2005, ApJ, 618, 569
Myers, A. T., McKee, C. F., Cunningham, A. J., Klein, R. I., & Krumholz, M. R. 2013, ApJ, 766, 97
Nagamine, K. 2010, Advances in Astronomy, 2010
Narayanan, D., & Hopkins, P. F. 2013, MNRAS, 433, 1223
Narayanan, D., et al. 2015, Nature, 525, 496
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Nelson, A. F. 2006, MNRAS, 373, 1039
Newman, S. F., et al. 2012, ApJ, 752, 111
Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., & Maeda, K. 2006, Nuclear Physics A, 777, 424
Novak, G. S., Ostriker, J. P., & Ciotti, L. 2011, ApJ, 737, 26
Nurmi, P., Heinämäki, P., Saar, E., Einasto, M., Holopainen, J., Martínez, V. J., & Einasto, J. 2006, A&A, in press [astro-ph/0611941]
Oñorbe, J., Boylan-Kolchin, M., Bullock, J. S., Hopkins, P. F., Kereš, D., Faucher-Giguère, C.-A., Quataert, E., & Murray, N. 2015, MNRAS, 454, 2092
Offner, S. S. R., Klein, R. I., McKee, C. F., & Krumholz, M. R. 2009, ApJ, 703, 131
Offner, S. S. R., Lee, E. J., Goodman, A. A., & Arce, H. 2011, ApJ, 743, 91
Oppenheimer, B. D., & Davé, R. 2006, MNRAS, 373, 1265
Oppenheimer, B. D., Davé, R., Kereš, D., Fardal, M., Katz, N., Kollmeier, J. A., & Weinberg, D. H. 2010, MNRAS, 406, 2325
O'Shea, B. W., Nagamine, K., Springel, V., Hernquist, L., & Norman, M. L. 2005, ApJS, 160, 1
Ostriker, E. C., & Shetty, R. 2011, ApJ, 731, 41
Padoan, P., Haugbølle, T., & Nordlund, Å. 2012, ApJL, 759, L27
Padoan, P., & Nordlund, Å. 2002, ApJ, 576, 870
—. 2011, ApJ, 730, 40
Pakmor, R., Springel, V., Bauer, A., Mocz, P., Munoz, D. J., Ohlmann, S. T., Schaal, K., & Zhu, C. 2016, MNRAS, 455, 1134
Pan, L., & Scannapieco, E. 2010, ApJ, 721, 1765
Pawlik, A. H., Milosavljević, M., & Bromm, V. 2013, ApJ, 767, 59
Petit, A. C., Krumholz, M. R., Goldbaum, N. J., & Forbes, J. C. 2015, MNRAS, 449, 2588
Pettini, M., Madau, P., Bolte, M., Prochaska, J. X., Ellison, S. L., & Fan, X. 2003, ApJ, 594, 695
Piontek, F., & Steinmetz, M. 2009, MNRAS, in press, arXiv:0909.4167
Planck Collaboration et al. 2013, ArXiv e-prints
Porter, D. H. 1985, PhD thesis, California Univ., Berkeley.
Powell, L. C., Slyz, A., & Devriendt, J. 2011, MNRAS, 414, 3671
Power, C., Navarro, J. F., Jenkins, A., Frenk, C. S., White, S. D. M., Springel, V., Stadel, J., & Quinn, T. 2003, MNRAS, 338, 14
Power, C., Read, J. I., & Hobbs, A. 2013, MNRAS, in press, arXiv:1307.0668
Price, D. J. 2008, Journal of Computational Physics, 227, 10040
—. 2012, Journal of Computational Physics, 231, 759
Price, D. J., & Federrath, C. 2010, MNRAS, 406, 1659
Price, D. J., & Monaghan, J. J. 2007, MNRAS, 374, 1347
Rafikov, R. R. 2007, ApJ, 662, 642
Rahmati, A., Schaye, J., Pawlik, A. H., & Raicevic, M. 2013, MNRAS, 431, 2261
Read, J. I., & Hayfield, T. 2012, MNRAS, 422, 3037
Read, J. I., Hayfield, T., & Agertz, O. 2010, MNRAS, 405, 1513
Richings, A. J., & Schaye, J. 2016, MNRAS, 458, 270
Ritchie, B. W., & Thomas, P. A. 2001, MNRAS, 323, 743
Robertson, B. E., & Kravtsov, A. V. 2008, ApJ, 680, 1083
Rodionov, S. A., & Sotnikova, N. Y. 2005, Astronomy Reports, 49, 470
Romeo, A. B. 1998, A&A, 335, 922
Rosdahl, J., Schaye, J., Dubois, Y., Kimm, T., & Teyssier, R. 2016, MNRAS, submitted, arXiv:1609.01296
Rosolowsky, E., Engargiola, G., Plambeck, R., & Blitz, L. 2003, ApJ, 599, 258
Rosswog, S. 2014, arXiv e-print: 1405.6034
Rybicki, G. B., & Lightman, A. P. 1986, Radiative Processes in Astrophysics (Wiley-VCH), 400
Saitoh, T. R., Daisaka, H., Kokubo, E., Makino, J., Okamoto, T., Tomisaka, K., Wada, K., & Yoshida, N. 2008, PASJ, 60, 667
Saitoh, T. R., & Makino, J. 2009, ApJL, 697, L99
—. 2013, ApJ, 768, 44
Sarazin, C. L. 1988, X-ray emission from clusters of galaxies (Cambridge: Cambridge University Press)
Sato, T., Martin, C. L., Noeske, K. G., Koo, D. C., & Lotz, J. M. 2009, ApJ, 696, 214
Schaye, J., et al. 2010, MNRAS, 402, 1536
Semenov, D., Henning, T., Helling, C., Ilgner, M., & Sedlmayr, E. 2003, A&A, 410, 611
Semenov, V. A., Kravtsov, A. V., & Gnedin, N. Y. 2015, ApJ, in press, arXiv:1512.03101
Shen, S., Madau, P., Conroy, C., Governato, F., & Mayer, L. 2014, ApJ, 792, 99
Shetty, R., & Ostriker, E. C. 2008, ApJ, 684, 978
Sijacki, D., Vogelsberger, M., Keres, D., Springel, V., & Hernquist, L. 2012, MNRAS, 424, 2999
Skinner, M. A., & Ostriker, E. C. 2015, ApJ, 809, 187
Slavin, J. D., & Cox, D. P. 1992, ApJ, 392, 131
Smagorinsky, J. 1963, Monthly Weather Review, 91, 99
Somerville, R. S., & Primack, J. R. 1999, MNRAS, 310, 1087
Songaila, A. 2005, AJ, 130, 1996
Sparre, M., Hayward, C. C., Feldmann, R., Faucher-Giguère, C.-A., Muratov, A. L., Kereš, D., & Hopkins, P. F. 2015, MNRAS, in press, arXiv:1510.03869
Spitzer, L., & Härm, R. 1953, Phys. Rev., 89, 977
Springel, V. 2005, MNRAS, 364, 1105
Springel, V. 2010, MNRAS, 401, 791
Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776
Springel, V., & Hernquist, L. 2003a, MNRAS, 339, 289
—. 2003b, MNRAS, 339, 312
Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726
Steidel, C. C., Erb, D. K., Shapley, A. E., Pettini, M., Reddy, N., Bogosavljević, M., Rudie, G. C., & Rakic, O. 2010, ApJ, 717, 289
Stinson, G., Seth, A., Katz, N., Wadsley, J., Governato, F., & Quinn, T. 2006, MNRAS, 373, 1074

Stinson, G. S., Brook, C., Macciò, A. V., Wadsley, J., Quinn, T. R., & Couchman, H. M. P. 2013, MNRAS, 428, 129
Stone, J. M., Gardiner, T. A., Teuben, P., Hawley, J. F., & Simon, J. B. 2008, ApJS, 178, 137
Su, K.-Y., Hopkins, P. F., Hayward, C. C., Faucher-Giguère, C.-A., Keres, D., Ma, X., & Robles, V. H. 2016, MNRAS, in press, arXiv:1607.05274
Tasker, E. J. 2011, ApJ, 730, 11
Tasker, E. J., & Tan, J. C. 2009, ApJ, 700, 358
Teyssier, R. 2002, A&A, 385, 337
—. 2015, ARA&A, 53, 325
Teyssier, R., Pontzen, A., Dubois, Y., & Read, J. I. 2013, MNRAS, 429, 3068
Thompson, T. A., Quataert, E., & Murray, N. 2005, ApJ, 630, 167
Thornton, K., Gaudlitz, M., Janka, H., & Steinmetz, M. 1998, ApJ, 500, 95
Torrey, P., Hopkins, P. F., Faucher-Giguère, C.-A., Vogelsberger, M., Quataert, E., Kereš, D., & Murray, N. 2016, MNRAS, in press, arXiv:1601.07186
Torrey, P., Vogelsberger, M., Sijacki, D., Springel, V., & Hernquist, L. 2012, MNRAS, 427, 2224
Truelove, J. K., Klein, R. I., McKee, C. F., Holliman, II, J. H., Howell, L. H., & Greenough, J. A. 1997, ApJL, 489, L179+
Uhlig, M., Pfrommer, C., Sharma, M., Nath, B. B., Enßlin, T. A., & Springel, V. 2012, MNRAS, 423, 2374
van de Voort, F., Quataert, E., Hopkins, P. F., Kereš, D., & Faucher-Giguère, C.-A. 2015, MNRAS, 447, 140
van den Hoek, L. B., & Groenewegen, M. A. T. 1997, A&AS, 123
Verner, D. A., & Ferland, G. J. 1996, ApJS, 103, 467
Vogelsberger, M., Sijacki, D., Keres, D., Springel, V., & Hernquist, L. 2012, MNRAS, 425, 3024
Vogelsberger, M., et al. 2014, Nature, 509, 177
Wadsley, J. W., Veeravalli, G., & Couchman, H. M. P. 2008, MNRAS, 387, 427
Walch, S., & Naab, T. 2015, MNRAS, 451, 2757
Wetzel, A. R., Hopkins, P. F., Kim, J.-h., Faucher-Giguère, C.-A., Kereš, D., & Quataert, E. 2016, ApJL, 827, L23
Wetzel, A. R., & White, M. 2010, MNRAS, 403, 1072
Wheeler, C., Oñorbe, J., Bullock, J. S., Boylan-Kolchin, M., Elbert, O. D., Garrison-Kimmel, S., Hopkins, P. F., & Kereš, D. 2015a, MNRAS, 453, 1305
Wheeler, C., Pace, A. B., Bullock, J. S., Boylan-Kolchin, M., Onorbe, J., Fitts, A., Hopkins, P. F., & Keres, D. 2015b, MNRAS, in press, arXiv:1511.01095
White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52
Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009a, MNRAS, 393, 99
Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., & Tornatore, L. 2009b, MNRAS, 399, 574
Williams, J. P., & McKee, C. F. 1997, ApJ, 476, 166
Wise, J. H., Abel, T., Turk, M. J., Norman, M. L., & Smith, B. D. 2012, MNRAS, 427, 311
Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, ApJ, 587, 278
Woosley, S. E., & Weaver, T. A. 1995, ApJS, 101, 181
Yan, H., & Lazarian, A. 2008, ApJ, 673, 942
Zel'dovich, Y. B. 1970, A&A, 5, 84
Zhu, Q., Hernquist, L., & Li, Y. 2014, ApJ, in press, arxiv:1410.4222
Zhu, Q., & Li, Y. 2016, ApJ, in press, arXiv:1606.00796
Zingale, M., et al. 2002, ApJS, 143, 539
Zuckerman, B., & Evans, II, N. J. 1974, ApJL, 192, L149

APPENDIX A: APPROXIMATE STELLAR EVOLUTION TABULATIONS

Here we present simple fits to the stellar evolution models and yields used in the FIRE simulations. We note that in several cases the simulations in this paper utilize a more detailed look-up table; however for all practical purposes the fits here are sufficiently accurate that the differences are negligible. Stellar evolution results are obtained from STARBURST99 (Leitherer et al. 1999) assuming a Kroupa (2001) IMF. SNe Ia rates follow Mannucci et al. (2006) including both prompt and delayed populations. Yields for core-collapse SNe are IMF-averaged for the same IMF, from the tables (including hypernovae) in Nomoto et al. (2006). Yields for SNe Ia follow Iwamoto et al. (1999). Yields for OB/AGB winds are taken from the synthesis of the models from van den Hoek & Groenewegen (1997), Marigo (2001), and Izzard et al. (2004) as synthesized in Wiersma et al. (2009b), appropriately re-averaged over the IMF.

(i) **SNe Ia:** These occur with rate-per-unit-stellar mass $R_{\rm Ia} = dN_{\rm Ia}/dt = 0$ for $t_{\rm Myr} < 37.53$ (where $t_{\rm Myr}$ is the age of the star particle in Myr), then $R_{\rm Ia}/({\rm SNe\,Myr^{-1}\,M_\odot^{-1}}) = 5.3\times10^{-8} + 1.6\times10^{-5}\,\exp\{-[(t_{\rm Myr}-50)/10]^2/2\}$ for $t_{\rm Myr} \geq 37.53$. Since the rate is per-stellar-mass, the expectation value of the number of SNe for a star particle of mass m_b and timestep Δt_b is $= R_{\rm Ia}(t^b_{\rm Myr})\,m_b\,\Delta t_b$, and the trial for "success" (an explosion) is drawn from a binomial distribution. Each SNe Ia has ejecta mass $M_{\rm ej} = 1.4\,M_\odot$ and energy $E_{\rm ej} = (1/2)\,M_{\rm ej}\,v_{\rm ej}^2 = 10^{51}$ erg. The ejecta yield mass for the tracked species is: (Z, He, C, N, O, Ne, Mg, Si, S, Ca, Fe) = (1.4, 0, 0.049, 1.2×10^{-6}, 0.143, 0.0045, 0.0086, 0.156, 0.087, 0.012, 0.743)\,$M_\odot$.

(ii) **SNe II:** The core-collapse SNe rate can be surprisingly well-fit by a simple piecewise-constant function, $R_{\rm II} = 0$ for $t_{\rm Myr} < 3.401$ or $t_{\rm Myr} > 37.53$; $R_{\rm II}/({\rm SNe\,Myr^{-1}\,M_\odot^{-1}}) = 5.408\times10^{-4}$ for $3.4 < t_{\rm Myr} < 10.37$; and $R_{\rm II}/({\rm SNe\,Myr^{-1}\,M_\odot^{-1}}) = 2.516\times10^{-4}$ for $10.37 < t_{\rm Myr} < 37.53$. The IMF-averaged ejecta mass per explosion is $M_{\rm ej} = 10.5\,M_\odot$, with ejecta energy $E_{\rm ej} = (1/2)\,M_{\rm ej}\,v_{\rm ej}^2 = 10^{51}$ erg. These are normalized so that the ejecta mass and energy exactly match the integrated totals from STARBURST99; the time-averaged energy injection rate ($\langle R_{\rm II}\,E_{\rm ej}\rangle$) is within $\sim 10\%$ of the tabulated STARBURST99 rate at all times. The IMF-averaged yields are (He, C, N, O, Ne, Mg, Si, S, Ca, Fe) = (3.87, 0.133, 0.0479$\tilde{N}$, 1.17, 0.30, 0.0987, 0.0933, 0.0397, 0.00458, 0.0741)\,$M_\odot$, where $\tilde{N} = {\rm MAX}(Z/Z_\odot, 1.65)$ accounts for the strongly progenitor-metallicity dependent N yields. The total metal yield Z is given by summing the explicitly-followed species, with an additional $\sim 2\%$ added to account for un-tracked species ($Z = 1.02\sum Z_{i,\,\rm followed}$); the remaining ejecta is H.[23]

(iii) **OB/AGB Mass-Loss:** We include all non-SNe mass-loss channels here, but this is dominated by OB/AGB-star winds. The

[23] For reference, the yields from Woosley & Weaver (1995), used in FIRE-1, are (He, C, N, O, Ne, Mg, Si, S, Ca, Fe) = (4.03, 0.117, 0.0399$\tilde{N}$, 1.06, 0.169, 0.0596, 0.0924, 0.0408, 0.00492, 0.0842)\,$M_\odot$. In both FIRE-1 and FIRE-2, we choose to omit the progenitor stellar metallicity dependence of the predicted yields (using the yields for solar-metallicity progenitors instead) if the standard deviation of the metallicity dependence $|dM_Z({\rm species})/dZ_{\rm progenitor}|$ between the models of Chieffi & Limongi (2004), Woosley & Weaver (1995), and Nomoto et al. (2006) is larger than the magnitude of the actual predicted trend $|dM_Z({\rm species})/dZ_{\rm progenitor}|$ from Nomoto et al. 2006. The only tracked species which passes this test is N. We do set the progenitor metallicity as a "floor" to the yields, following exactly the algorithm in Wiersma et al. (2009b).

IMF-integrated mass-loss rate for a stellar population/particle of mass M_* is $\dot{M}_{\rm w} = M_*\,f_{\rm w}\,{\rm Gyr}^{-1}$ with $f_{\rm w} = 4.763\,(0.01+Z/Z_\odot)$ for $t_{\rm Myr}<1$; $f_{\rm w} = 4.763\,(0.01+Z/Z_\odot)\,t_{\rm Myr}^{1.45+0.8\,\ln(Z/Z_\odot)}$ for $1<t_{\rm Myr}<3.5$; $f_{\rm w} = 29.4\,(t_{\rm Myr}/3.5)^{-3.25}+0.0042$ for $3.5<t_{\rm Myr}<100$; and $f_{\rm w} = 0.42\,(t_{\rm Myr}/1000)^{-1.1}/(19.81-\ln(t_{\rm Myr}))$ for $t_{\rm Myr}>100$. The total (IMF-averaged) kinetic luminosity of the mass-loss is given by $L_{\rm kinetic} = (1/2)\,\dot{M}_{\rm w}\,\langle v_{\rm w}^2\rangle = \psi\times10^{12}\,{\rm erg\,g^{-1}}\,\dot{M}_{\rm w}$, with $\psi = (5.94\times10^4)/(1+(t_{\rm Myr}/2.5)^{1.4}+(t_{\rm Myr}/10)^{5.0})+4.83$ for $t_{\rm Myr}<100$ and $\psi=4.83$ for $t_{\rm Myr}>100$. The yields are given by the maximum of either the progenitor stellar surface abundances or, for the light species (He, C, N, O), mass fractions $=(0.36, 0.016, 0.0041\,\tilde{N}, 0.0118)$.

(iv) **Radiation:** We define the light-to-mass ratio in a given band $\Psi_{\rm band}$, with units $L_\odot/M_\odot$. Then the bolometric $\Psi_{\rm bol}=1136.59$ for $t_{\rm Myr}<3.5$, and $\Psi_{\rm bol} = 1500\exp[-4.145\,x+0.691\,x^2-0.0576\,x^3]$ with $x\equiv\log_{10}(t_{\rm Myr}/3.5)$ for $t_{\rm Myr}>3.5$. For the bands used in our radiation hydrodynamics, we have the following intrinsic (before attenuation) bolometric corrections. In the mid/far IR, $\Psi_{\rm IR}=0$. In optical/NIR, $\Psi_{\rm opt} = f_{\rm opt}\,\Psi_{\rm bol}$ with $f_{\rm opt}=0.09$ for $t_{\rm Myr}<2.5$; $f_{\rm Opt} = 0.09\,(1+[(t_{\rm Myr}-2.5)/4]^2)$ for $2.5<t_{\rm Myr}<6$; $f_{\rm Opt} = 1-0.841/(1+[(t_{\rm Myr}-6)/300])$ for $t_{\rm Myr}>6$. For the photo-electric FUV band $\Psi_{\rm FUV} = 271\,[1+(t_{\rm Myr}/3.4)^2]$ for $t_{\rm Myr}<3.4$; $\Psi_{\rm FUV} = 572\,(t_{\rm Myr}/3.4)^{-1.5}$ for $t_{\rm Myr}>3.4$. For the ionizing band $\Psi_{\rm ion}=500$ for $t_{\rm Myr}<3.5$; $\Psi_{\rm ion} = 60\,(t_{\rm Myr}/3.5)^{-3.6}+470\,(t_{\rm Myr}/3.5)^{0.045-1.82\,\ln t_{\rm Myr}}$ for $3.5<t_{\rm Myr}<25$; $\Psi_{\rm ion}=0$ for $t_{\rm Myr}>25$. The remaining UV luminosity, $\Psi_{\rm bol}-(\Psi_{\rm IR}+\Psi_{\rm opt}+\Psi_{\rm FUV}+\Psi_{\rm ion})$ is assigned to the NUV band $\Psi_{\rm NUV}$. The flux-mean dust opacities adopted are $(\kappa_{\rm FUV},\,\kappa_{\rm NUV},\,\kappa_{\rm opt},\,\kappa_{\rm IR}) = (2000, 1800, 180, 10)\,(Z/Z_\odot)\,{\rm cm^2\,g^{-1}}$. The photo-ionization rate (and corresponding $\kappa_{\rm ion}$) is calculated from the neutral hydrogen density as described in Appendix E and Appendix B below.

APPENDIX B: APPROXIMATE COOLING FUNCTIONS

Here we provide simple fitting-function approximations to the complete set of cooling functions used in our FIRE simulations. Note that for several of these, we use somewhat more accurate look-up tables in the simulations (as a function of temperature, density, and metallicity), but we provide functions accurate to $\sim10\%$ over the relevant dynamic range in the simulations ($\sim10-10^9$ K), so that interested readers can reproduce our full cooling physics treatment.

The instantaneous cooling+heating rate per unit volume is given by the sum over all processes,

$$\frac{de^a_{\rm thermal}}{dt} = -n_{{\rm H},a}^2\,\Lambda^a_{\rm net} = -n_{{\rm H},a}^2\sum_i\Lambda_i \tag{B1}$$

where $e^a_{\rm thermal}$ is the thermal energy density of a gas element a. Here we will use Λ to denote both cooling and heating rates, but with opposite signs; a positive sign here denotes cooling, and $\tilde{n}_x\equiv n_x/n_{\rm H}$ is the number of species x per hydrogen nucleus (e.g. $\tilde{n}_e$, $\tilde{n}_{\rm H0}$, and $\tilde{n}_{\rm HeI}$ denote the free electron, neutral hydrogen, and neutral helium numbers). Below, all units are cgs ($[T]={\rm K}$, $[n_{\rm H}]={\rm cm^{-3}}$, $[\Lambda] = {\rm erg\,s^{-1}\,cm^3}$). The processes we track include:

(i) **Collisional Excitation**: from Katz et al. (1996):

$$\Lambda_{\rm CE} = (\beta_{\rm H0}\,\tilde{n}_{\rm H0}+\beta_{\rm HeI}\,\tilde{n}_{\rm HeI})\,\tilde{n}_e \tag{B2}$$

$$\beta_{\rm H0} = 7.50\times10^{-19}\,\tau_5\,\exp\left(-\frac{118348}{T}\right) \tag{B3}$$

$$\beta_{\rm HeI} = 5.54\times10^{-17}\,\tau_5\,T^{-0.397}\,\exp\left(-\frac{473638}{T}\right) \tag{B4}$$

$$\tau_5 \equiv \left[1+\left(\frac{T}{10^5}\right)^{1/2}\right]^{-1} \tag{B5}$$

(ii) **Collisional Ionization**: also from Katz et al. (1996):

$$\Lambda_{\rm CI} = 10^{-11}\,\tilde{n}_e\times (2.18\,\gamma_{\rm H0}\,\tilde{n}_{\rm H0}+3.94\,\gamma_{\rm He0}\,\tilde{n}_{\rm He0}+8.72\,\gamma_{\rm HeI}\,\tilde{n}_{\rm HeI}) \tag{B6}$$

$$\gamma_{\rm H0} = 5.85\times10^{-11}\,T^{1/2}\,\tau_5\,\exp\left(-\frac{157809.1}{T}\right) \tag{B7}$$

$$\gamma_{\rm He0} = 2.38\times10^{-11}\,T^{1/2}\,\tau_5\,\exp\left(-\frac{285335.4}{T}\right) \tag{B8}$$

$$\gamma_{\rm HeI} = 5.68\times10^{-12}\,T^{1/2}\,\tau_5\,\exp\left(-\frac{631515.0}{T}\right) \tag{B9}$$

(iii) **Recombination**: from Verner & Ferland (1996):

$$\Lambda_{\rm Rec} = 1.036\times10^{-16}\,T\,\tilde{n}_e\times\left(\alpha_{\rm HI}\,\tilde{n}_{\rm HI}+\left[\alpha_{\rm HeI}+\frac{629922.78}{T}\,\alpha_{\rm di}\right]\tilde{n}_{\rm HeI}+\alpha_{\rm HeII}\,\tilde{n}_{\rm HeII}\right) \tag{B10}$$

$$\alpha_{\rm HI} = 7.982\times10^{-11}\left(\frac{1.774}{T^{0.5}}\right)\times\left(1+\frac{T^{0.5}}{1.774}\right)^{-0.252}\left(1+\frac{T^{0.5}}{838.81}\right)^{-1.748} \tag{B11}$$

$$\alpha_{\rm HeI} = 9.356\times10^{-10}\left(\frac{0.2065}{T^{0.5}}\right)\times\left(1+\frac{T^{0.5}}{0.2065}\right)^{-0.2108}\left(1+\frac{T^{0.5}}{6063.0}\right)^{-1.7892} \tag{B12}$$

$$\alpha_{\rm HeII} = 1.5964\times10^{-10}\left(\frac{2.5092}{T^{0.5}}\right)\times\left(1+\frac{T^{0.5}}{2.5092}\right)^{-0.252}\left(1+\frac{T^{0.5}}{1677.6}\right)^{-1.748} \tag{B13}$$

$$\alpha_{\rm di} = 1.9\times10^{-3}\,T^{1.5}\,\exp\left(-\frac{470000}{T}\right)\times\left(1+0.3\,\exp\left[-\frac{94000}{T}\right]\right) \tag{B14}$$

Here the $\alpha_{\rm di}$ term comes from dielectric recombination.

(iv) **Free-free emission**: from Rybicki & Lightman (1986):

$$\Lambda_{\rm FF} = \beta_{\rm ff}(T)\,(\tilde{n}_{\rm HI}+\tilde{n}_{\rm HeI}+4\,\tilde{n}_{\rm HeII})\,\tilde{n}_e \tag{B15}$$

$$\beta_{\rm ff}(T) = 1.43\times10^{-27}\,T^{1/2}\times\left[1.1+0.34\,\exp\left\{-(5.5-\log_{10}[T])^2/3\right\}\right] \tag{B16}$$

(v) **High-Temperature Metal-Line Cooling**: this refers to metal-line cooling processes in mostly ionized gas, with temperatures $\gtrsim10^4$ K. We use the public look-up tables from Wiersma et al. (2009a), for which:

$$\Lambda_{\rm Metal} = \sum_{\rm species\ }{}_i\,\Lambda^i_{\rm Metal} = \sum_i\tilde{n}_e\,\xi_i(n_{\rm H},T,z)\,\frac{Z^i}{Z^i_\odot} \tag{B17}$$

where this refers to the sum over all tracked metal species i (here

C, N, O, Ne, Mg, Si, S, Ca, Fe), and $Z^i/Z^i_\odot$ is the abundance of species i relative to solar. We adopt solar abundances (Z, He, C, N, O, Ne, Mg, Si, S, Ca, Fe) = (0.02, 0.28, 3.26×10^{-3}, 1.32×10^{-3}, 8.65×10^{-3}, 2.22×10^{-3}, 9.31×10^{-4}, 1.08×10^{-3}, 6.44×10^{-4}, 1.01×10^{-4}, 1.73×10^{-3}), which give abundance ratios matching Asplund et al. (2009) scaled to the total metallicity = 0.02 (used because it matches the assumed values in the cooling computations and stellar evolution models). The functions ξ_i depend on density, temperature, and redshift z (because they assumes photo-ionization by a redshift-dependent UV background); they are taken from the look-up tables provided by Wiersma et al. (2009a).[24]

(vi) **Low-Temperature Metal Line, Fine-Structure, & Molecular Cooling**: this combines the gas-phase low-temperature cooling (including molecular and atomic processes) in mostly neutral gas below $\lesssim 10^4$ K. From our compilation of CLOUDY runs (Ferland et al. 1998), fitting the resulting look-up tables, we obtain approximately:

$$\Lambda_{\rm Cold} = 2.896\times10^{-26}\left\{\left(\frac{T}{125.215}\right)^{-4.9202} + \left(\frac{T}{1349.86}\right)^{-1.7288} + \left(\frac{T}{6450.06}\right)^{-0.3075}\right\}^{-1} \times \left(\frac{1+(Z/Z_\odot)}{1+0.00143\,n_{\rm H}}\right)(1-f_{\rm selfshield}) \times \left(0.001 + \frac{0.10\,n_{\rm H}}{1+n_{\rm H}} + \frac{0.09\,n_{\rm H}}{1+0.1\,n_{\rm H}} + \frac{(Z/Z_\odot)^2}{1+n_{\rm H}}\right) \times \exp\left(-\left[\frac{T}{158000}\right]^2\right) \quad \text{(B18)}$$

where $f_{\rm selfshield}$ accounts for the local radiation environment by applying a simple (fitted) local shielding correction for UV/ionizing photons, $f_{\rm selfshield} \equiv \exp(-\tilde{\tau}_a^{\rm ion})$ with $\tilde{\tau}_a^{\rm ion} \equiv \sigma^H_{\nu_0}\, n_{{\rm H},a}\, \ell_a^{\rm fit}$ where $\sigma^H_{\nu_0} \equiv 6\times10^{-18}\,{\rm cm}^{-2}$ and $\ell_a^{\rm fit} \equiv 4.4\,{\rm pc}\,(T/10^4\,K)^{-0.173}\,\Gamma_{-12}^{-2/3}$ (Γ_{-12} is the ionization rate in units of $10^{-12}\,{\rm s}^{-1}$, including both the UV background and local sources assuming they have the same spectral shape, as defined below for photo-ionization heating).

(vii) **Dust Collisional Heating/Cooling**: from Meijerink & Spaans (2005):

$$\Lambda_{\rm Dust} = 1.12\times10^{-32}\,(T-T_{\rm dust})\,T^{1/2} \times \left(1-0.8\exp\left[-\frac{75}{T}\right]\right)\left(\frac{Z}{Z_\odot}\right) \quad \text{(B19)}$$

where we take $T_{\rm dust} = 30\,{\rm K}$ here, and the $Z/Z_\odot$ term comes from assuming a constant dust-to-metals ratio.

(viii) **Compton Heating/Cooling**: from the CMB, gives (Rybicki & Lightman 1986):

$$\Lambda_{\rm Compton} = 5.65\times10^{-36}\,\tilde{n}_e\,(T-T_{\rm CMB}[z])\,(1+z)^4\,n_{\rm H}^{-1} \quad \text{(B20)}$$

(ix) **Photo-Ionization Heating**: from the UVB and local (in-simulation stellar sources) gives a heating rate, hence negative Λ, of

$$\Lambda_{\rm ion} = -\tilde{f}\,(\epsilon_{\rm H0}\,\tilde{n}_{\rm H0} + \epsilon_{\rm He0}\,\tilde{n}_{\rm He0} + \epsilon_{\rm HeI}\,\tilde{n}_{\rm HeI})\,n_{\rm H}^{-1} \quad \text{(B21)}$$

$$\tilde{f} \equiv \left(1+\frac{e^{\rm local}_{\nu,\,\rm ion}}{e^{\rm UVB}_{\nu,\,\rm ion}}\right) f_{\rm selfshield} \quad \text{(B22)}$$

$$\log_{10}(\epsilon_{\rm H0}) \approx -24.6 + 1.62\,x + 14.9\,x^2 - 45.5\,x^3 + 46.2\,x^4 - 16.7\,x^5 - \exp[50\,(x-1.05)] \quad \text{(B23)}$$

$$\log_{10}(\epsilon_{\rm He0}) \approx \log_{10}(\epsilon_{\rm H0}) - 0.0366 + 0.376\,x$$

$$\log_{10}(\epsilon_{\rm HeI}) \approx -26.3 + 0.816\,x + 78.2\,x^2 - 837\,x^3 + 4770\,x^4 - 15600\,x^5 + 29600\,x^6 - 32400\,x^7 + 18900\,x^8 - 4550\,x^9$$

$$x \equiv \log_{10}(1+z)$$

where $\epsilon_{\rm H0}$, $\epsilon_{\rm He0}$, and $\epsilon_{\rm HeI}$ are pre-tabulated for the assumed UV background magnitude and shape, as a function of redshift in Faucher-Giguère et al. (2009)[25] – values above are simple polynomial fits good to $\sim 10\%$ up to $z\sim10$. The factor $\tilde{f}$ accounts for both self-shielding (reducing the effective incident flux by $f_{\rm selfshield}$) and the contribution from local sources, where $e^{\rm local}_{\nu,\,\rm ion}$ is the ionizing band radiation energy density calculated explicitly from the radiation-hydrodynamic treatment in the code (Appendix E), and $e^{\rm UVB}_{\nu,\,\rm ion}$ is the meta-galactic UV background (UVB) energy density integrated in the same band for the same wavelength range (H-ionizing frequencies). Note this means we assume the spectral slope of escaping, ionizing radiation from resolved stars in the simulation is the same as the UVB.

(x) **Cosmic Ray Heating**: from Guo & Oh (2008):

$$\Lambda_{\rm CR} = -1.0\times10^{-16}\,(0.98+1.65\,\tilde{n}_e\,X_H)\,e_{\rm CR}\,n_{\rm H}^{-1} \quad \text{(B24)}$$

where we assume a uniform MW-like cosmic ray background, $e_{\rm CR}\approx9.0\times10^{-12}$ for densities $n_{\rm H}>0.01$, and at lower densities $e_{\rm CR} = 9.0\times10^{-12}\,(n_{\rm H}/0.01)$ (this is just to avoid predicting a much higher CR heating rate in some regions only because they have extremely low densities).

(xi) **Photo-Electric Heating**: from Wolfire et al. (2003):

$$\Lambda_{\rm PE} = -1.3\times10^{-24}\,\tilde{e}^{\rm pe}_\nu\,n_{\rm H}^{-1}\left(\frac{Z}{Z_\odot}\right) \times \left(\frac{0.049}{1+(x_{\rm pe}/1925)^{0.73}} + \frac{0.037\,(T/10^4)^{0.7}}{1+(x_{\rm pe}/5000)}\right) \quad \text{(B25)}$$

$$x_{\rm pe} \equiv \frac{\tilde{e}^{\rm pe}_\nu\,T^{0.5}}{0.5\,\tilde{n}_e\,n_{\rm H}} \quad \text{(B26)}$$

where $\tilde{e}^{\rm pe}_\nu$ is the photon energy density in the photo-electric band, normalized to the Habing (1968) MW units, $\tilde{e}^{\rm pe}_\nu \equiv e^{\rm pe}_\nu/(3.9\times10^{-14}\,{\rm erg\,cm^{-3}})$. The $Z/Z_\odot$ term comes from assuming a constant dust-to-metals ratio. Here the field $e^{\rm pe}_\nu$ is the FUV band radiation energy density calculated explicitly from the radiation-hydrodynamic treatment in the code, described in Appendix E.

(xii) **Magneto-Hydrodynamic Work & Shocks**: from the MHD equations, we obtain some fluid-dynamic change to the temperature (owing to compression/expansion, shocks, etc). We include this self-consistently in the fully-implicit temperature update:

$$\Lambda_{\rm MHD} = -\mu\left.\frac{\partial u_{\rm thermal}}{\partial t}\right|_{\rm MHD} n_{\rm H}^{-1} \quad \text{(B27)}$$

where $u_{\rm thermal}$ is the specific internal energy (internal energy per unit mass).

[24] Available at http://www.strw.leidenuniv.nl/WSS08

[25] See http://galaxies.northwestern.edu/uvb

(xiii) **Optically-Thick Cooling**: lacking a full radiative transfer solution for cooling radiation, we approximate the effects of optically-thick cooling using the method from Rafikov (2007), which captures the most important effects by approximating each element as a "slab" with column density estimated via the Sobolev approximation and integrating a vertical atmosphere through to its photosphere to determine the net photon escape. This amounts to first summing the contributions above to determine the net heating/cooling rate $\Lambda_{\rm Net}$, and then restricting this to the cooling rate of said slab:

$$|\Lambda_{\rm Net}| < \Lambda_{\rm BB} \tag{B28}$$

$$\Lambda_{\rm BB} \equiv 5.67\times10^{-5}\,T^{4}\left(\frac{\mu}{\Sigma_{\rm eff}}\right)\frac{1}{1+\kappa_{\rm eff}\,\Sigma_{\rm eff}}\,n_{\rm H}^{-1} \tag{B29}$$

where $\Sigma_{\rm eff} = \langle \Sigma^{a,\,\rm Sobolev}_{\rm column}\rangle_{\phi,\theta} = \rho_a\,(h_a + \rho_a/|\nabla\rho_a|)$ uses the local Sobolev approximation to estimate the column density to infinity and is defined in the text (Appendix E), and $\kappa_{\rm eff}$ is the effective opacity.[26]

As noted in the text, the actual heating/cooling step is solved fully implicitly for each element on its own timestep.

APPENDIX C: ALGORITHMIC IMPLEMENTATION OF STAR FORMATION

(i) **Self-Gravitating:** First, following standard sink-particle approaches, we calculate the virial parameter α (ratio of thermal plus kinetic energy to potential energy) in a resolution element and allow only star formation in bound particles with $\alpha < 1$. From Hopkins et al. (2013d),

$$\alpha \equiv \frac{\|\nabla\otimes\mathbf{v}\|_a^2 + (c_{s,a}/h_a)^2}{8\pi\,G\,\rho_a} \tag{C1}$$

where $\|\nabla\otimes\mathbf{v}\|_a$ the Frobenius norm ($\|\mathbf{A}\|^2 \equiv \sum_{\alpha\beta\gamma...} A^2_{\alpha\beta\gamma...}$) of the velocity gradient tensor ($\otimes$ is the outer product), $c_{s,a}$ is the sound speed, ρ_a the density, and h_a the usual resolution scale (inter-particle spacing). This has the advantage that it converges to an explicitly resolution-independent expression in the super-sonic turbulence limit (see Hopkins et al. 2013d). Note that in GIZMO, we always use the higher-order accurate, matrix-based gradient estimators described in Hopkins (2015), which remain second-order accurate, consistent, and robust despite arbitrary particle configurations within the stencil (see also Maron & Howes 2003; Luo et al. 2008; Lanson & Vila 2008; Mocz et al. 2014; Pakmor et al. 2016). Using the zeroth-order inaccurate SPH gradient estimator, in contrast, gives similar results statistically, but makes identification of individual physically collapsing clouds much more noisy.

(ii) **Self-Shielding:** If $\alpha < 1$, we next calculate the shielded/molecular fraction $f_{\rm shielded}$, which is the fraction of the mass that should be able to self-shield and so cool efficiently (hence fragment to stellar mass scales). The expression for the shielded fraction from Krumholz & Gnedin (2011) is:

$$f_{\rm shielded} \equiv 1 - \frac{3}{1+4\,\tilde{\psi}_a} \tag{C2}$$

$$\tilde{\psi}_a \equiv \frac{0.6\,\tilde{\tau}_a\,(0.01+Z_a/Z_\odot)}{\ln\left(1+0.6\,\tilde{\phi}_a+0.01\,\tilde{\phi}_a^2\right)} \tag{C3}$$

$$\tilde{\phi}_a \equiv 0.756\,(1+3.1\,Z_a/Z_\odot)^{0.365} \tag{C4}$$

$$\tilde{\tau}_a \equiv 434.8\,{\rm cm^2\,g^{-1}}\,\rho_a\left(h_a + \frac{\rho_a}{|\nabla\rho|_a}\right) \tag{C5}$$

We require $f_{\rm shielded} > 0$ for star formation.

(iii) **Jeans-Unstable:** If $\alpha < 1$, $f_{\rm shielded} > 0$, we calculate the Jeans mass m_J, and only allow star formation in Jeans-unstable particles, specifically those where $m_J < m_{J,\rm crit} \equiv {\rm MAX}(m_a,\, 10^3\,M_\odot)$, where m_a is the particle mass. We calculate the Jeans mass as

$$m_J = 2\,M_\odot\left(\frac{c_{s,a}}{0.2\,{\rm km\,s^{-1}}}\right)^3\left(\frac{n_a}{10^3\,{\rm cm^{-3}}}\right)^{-1/2} \tag{C6}$$

where $n_a \equiv \rho_a/\mu_a$ is the gas number density.

(iv) **Sufficiently-Dense:** If $\alpha < 1$, $f_{\rm shielded} > 0$, $m_J < m_{J,\rm crit}$, we check if $n_a > n_{\rm crit}$, where $n_{\rm crit} = 1000\,{\rm cm^{-3}}$ is a minimum density (and $n_a \equiv n_{a,H} = X_H\,\rho_a/m_p$), to prevent spurious triggering of the above criteria in low-density gas.

We then assign the gas particle a volume-integrated SFR:

$$\dot{m}_*^a = \Theta(\alpha_a,\, f_{\rm shielded,a},\, n_a,\, m_{J,a})\,\frac{f_{\rm shielded}\,m_{{\rm gas},a}}{t_{{\rm freefall},a}} \tag{C7}$$

$$t_{{\rm freefall},a} \equiv \sqrt{\frac{3\pi}{32\,G\,\rho_a}} \tag{C8}$$

$$\Theta = \begin{cases} 1 & (\alpha<1,\, f_{\rm shielded}>0,\, m_J < m_{J,\rm crit},\, n_a > n_{\rm crit}) \\ 0 & \text{otherwise} \end{cases} \tag{C9}$$

Because we wish to maintain equal stellar and gas element masses, at each timestep Δt_a we assign the gas particle a probability $p_a = 1-\exp\left(-\dot{m}_*^a\,\Delta t_a/m_{\rm gas}^a\right)$ of turning into a star particle that timestep; we draw a uniform random variable $0<x<1$ and if $x < p_a$, we convert the gas particle to a star particle. It inherits all relevant properties of its parent particle.

APPENDIX D: ALGORITHMIC IMPLEMENTATION OF MECHANICAL FEEDBACK

Here we describe our implementation of mechanical feedback, used for SNe (Types Ia & II) and stellar mass-loss. This algorithm was first developed and presented in a series of papers, beginning with Hopkins et al. (2012c), and the version used in FIRE-1 (which contains most of the important features here up to some specific numerical improvements for FIRE-2) was presented in detail in Hopkins et al. (2014). Parts of that algorithm – in particular the treatment of SNe momentum accounting for the terminal momentum – were subsequently ported into other codes by Kimm & Cen (2014), Martizzi et al. (2015), and Rosdahl et al. (2016).

In Paper II, we discuss each piece of this algorithm in detail, and consider a large suite of both idealized test problems and full cosmological simulations, to test and validate each and show exactly how it does or does not influence our predictions. However for the sake of completeness and for those readers interested in reproducing our results, we include the complete algorithm here.

(i) Every timestep Δt_a, for each star particle a (at position $\mathbf{x}_a$), we first determine whether an "event" occurs: a SNe Ia, SNe II,

[26] We take $\kappa_{\rm eff}$ for dust ($T < 1500$) from the detailed tables in Semenov et al. (2003), assuming the dust, gas, and radiative temperatures are in equilibrium (true in the optically thick limit at these temperatures), which is approximately well-fit by $\kappa_{\rm eff} = 5$ for $150 \le T \le 1500$ and $\kappa_{\rm eff} = 0.0027\,T^{3/2}$ for $T < 150$. At higher temperatures the system is rarely optically thick, but for completeness we compute $\kappa_{\rm eff}$ from the gas-phase using standard approximations for stellar atmospheres: $\kappa_{\rm eff}^{-1} = \kappa_{\rm rad}^{-1} + \kappa_{\rm cond}^{-1}$, with $\kappa_{\rm cond} = 2.6\times10^{-7}\,\tilde{n}_e\,T^2\,\rho^{-2}$, $\kappa_{\rm rad} = \kappa_{\rm mol} + 1/(\kappa_{H^-}^{-1} + [\kappa_e+\kappa_{\rm Kr}]^{-1})$, $\kappa_{\rm mol} = 0.1\,Z$, $\kappa_e = 0.2\,(1+X_H)$, $\kappa_{H^-} = 1.1\times10^{-25}\,(Z\,\rho)^{1/2}\,T^{7.7}$, $\kappa_{\rm Kr} = 4.0\times10^{25}\,(1+X_H)\,Z\,\rho^{3.5}$.

and/or non-zero stellar mass-loss. This follows the rates and algorithms in Appendix A. If an event occurs, each gives some initial ejecta (or wind) mass $m_{\rm ej}$, metal mass $m_{Z,\rm ej}$ (defined for each species we track), momentum $p_{\rm ej} = m_{\rm ej}\,v_{\rm ej}$, and energy $E_{\rm ej}$. These also are given in Appendix A (for winds, $m_{\rm ej} = \Delta t\,\dot{M}_w$ from the star, for SNe it is the ejecta mass).

(ii) Identify gas elements surrounding the star particle: in a grid code this is straightforward, but in our mesh-free method, we define an effective neighbor number $N_{\rm ngb} = (4\pi/3)\,H_a^3\,\bar{n}_a(H_a)$ in the same manner as for hydrodynamics, where W is the kernel function, $\bar{n}_a = \sum W(\mathbf{x}_{ba} \equiv \mathbf{x}_b - \mathbf{x}_a, H_a)$, and H_a is the search radius. The equation for $N_{\rm ngb}(H_a)$ is non-linear so is solved iteratively in the neighbor search; see Springel (2005). Thus we obtain all gas elements b within a radius $|\mathbf{x}_{ba}| < H_a$ (where "a sees b"); we *also* identify all neighbors with $|\mathbf{x}_{ba}| < H_b$ (i.e. "b sees a"). We show in Paper II that this is important to ensure the "effective faces" close and the resulting distribution of ejecta is isotropic, in regions with highly disordered gas element positions.

(iii) Assign "vector weights" to each neighbor, by first boosting to the rest-frame of the star ($\mathbf{x}_a = \mathbf{0}$, $\mathbf{v}_a \equiv d\mathbf{x}_a/dt = \mathbf{0}$; in which the ejecta should be isotropic), then calculating the "effective face" that would be seen by the star particle (using the *same* definitions of inter-cell faces used in the hydrodynamics) and integrating the ejecta over solid angle through to each face. This amounts to defining the vector weight function $\mathbf{w}_{ba}$

$$\bar{\mathbf{w}}_{ba} \equiv \frac{\mathbf{w}_b}{\sum_c |\mathbf{w}_{ca}|} \tag{D1}$$

$$\mathbf{w}_{ba} \equiv \omega_{ba} \sum_{+,-} \sum_{\alpha} (\hat{\mathbf{x}}_{ba}^{\pm})^{\alpha}\,\left(f_{\pm}^{\alpha}\right)_a \tag{D2}$$

$$\left(f_{\pm}^{\alpha}\right)_a \equiv \left\{ \frac{1}{2}\left[1 + \left(\frac{\sum_c \omega_{ca}\,|\hat{\mathbf{x}}_{ca}^{\mp}|^{\alpha}}{\sum_c \omega_{ca}\,|\hat{\mathbf{x}}_{ca}^{\pm}|^{\alpha}}\right)^2\right]\right\}^{1/2} \tag{D3}$$

$$\omega_{ba} = \frac{\Delta\Omega_{ba}}{4\pi} \equiv \frac{1}{2}\left(1 - \frac{1}{\sqrt{1 + (\mathbf{A}_{ba}\cdot\hat{\mathbf{x}}_{ba})/(\pi\,|\mathbf{x}_{ba}|^2)}}\right) \tag{D4}$$

where $\mathbf{A}_{ba}$ is the effective vector face between elements b and a used in the finite-volume hydrodynamic calculations,[27] and the $\hat{\mathbf{x}}_{ca}^{\pm}$ are the positive or negative (singly-signed) projection vectors:

$$\hat{\mathbf{x}}_{ba} \equiv \frac{\mathbf{x}_{ba}}{|\mathbf{x}_{ba}|} = \sum_{+,-} \hat{\mathbf{x}}_{ba}^{\pm} \tag{D8}$$

$$(\hat{\mathbf{x}}_{ba}^{+})^{\alpha} \equiv |\mathbf{x}_{ba}|^{-1}\,\mathrm{MAX}(\mathbf{x}_{ba}^{\alpha}, 0)\Big|_{\alpha=x,y,z} \tag{D9}$$

$$(\hat{\mathbf{x}}_{ba}^{-})^{\alpha} \equiv |\mathbf{x}_{ba}|^{-1}\,\mathrm{MIN}(\mathbf{x}_{ba}^{\alpha}, 0)\Big|_{\alpha=x,y,z} \tag{D10}$$

These expressions are complicated but are derived in detail in Paper II. Their important properties are (1) they maintain *manifest* conservation of mass, momentum, and energy (see below). (2) They give fluxes which are statistically isotropic in the rest frame of the star, i.e. the ejecta are not systematically biased in one direction or another, even if there is a global density gradient such that there are, on average, more gas elements in one direction. We demonstrate this in numerical tests explicitly in Paper II, and show that many simpler prescriptions lead to systematic, unphysical biases in the ejecta deposition, e.g. if there is a thin, dense disk such that more gas neighbors are "in the disk," simple weighting $\bar{\mathbf{w}}_{ba}$ proportional to, say, the SPH kernel, leads to almost all the ejecta being coupled in the disk, driving an expanding ring, with almost no ejecta going into the vertical direction – when in fact the converged solution to this problem is exactly the opposite (hot gas "vents" in the vertical direction). (3) They approximate, as closely as possible without an expensive numerical quadrature, the exact integral of the ejecta through and into the "domains" of each gas neighbor determined by the hydrodynamic volume partition.

(iv) Assign initial fluxes in the rest-frame of the star:

$$\Delta m_b = |\bar{\mathbf{w}}_b|\,m_{\rm ej} \tag{D11}$$

$$\Delta m_{Z,b} = |\bar{\mathbf{w}}_b|\,m_{Z,\rm ej} \tag{D12}$$

$$\Delta E_b = |\bar{\mathbf{w}}_b|\,E_{\rm ej} \tag{D13}$$

$$\Delta \mathbf{p}_b = \bar{\mathbf{w}}_b\,p_{\rm ej} \tag{D14}$$

It is easy to see that our definition of $\bar{\mathbf{w}}_{ba}$ guarantees *exact* conservation of mass, energy, and linear momentum, and that the correct total radial (outward) momentum is assigned, e.g.:

$$\sum \Delta m_b = m_{\rm ej} \tag{D15}$$

$$\sum \Delta m_{Z,b} = m_{Z,\rm ej} \tag{D16}$$

$$\sum \Delta E_b = E_{\rm ej} \tag{D17}$$

$$\sum |\Delta \mathbf{p}_b| = p_{\rm ej} \tag{D18}$$

$$\sum \Delta \mathbf{p}_b = \mathbf{0} \tag{D19}$$

(v) Boost back to the simulation ("lab") frame: if the star is moving with velocity $\mathbf{v}_a$, then this boost transforms the momentum and energy fluxes:

$$\Delta \mathbf{p}_b' \equiv \Delta \mathbf{p}_b + \Delta m_b\,\mathbf{v}_a \tag{D20}$$

$$\Delta E_b' \equiv \Delta E_b + \frac{1}{2\,\Delta m_b}\left(|\Delta \mathbf{p}_b'|^2 - |\Delta \mathbf{p}_b|^2\right) \tag{D21}$$

(the mass fluxes are unchanged, $\Delta m_b' = \Delta m_b$, $\Delta m_{Z,b}' = \Delta m_{Z,b}$). Of course this maintains manifest conservation: the total momentum added to the neighbors via the $\Delta m_b\,\mathbf{v}_a$ term exactly cancels that lost by the star, since its mass decreases by $\sum \Delta m_b = m_{\rm ej}$.

(vi) Account for *PdV* (mechanical) work: consider that we have a particle b representing a volume domain with mass m_b around our source, with some mean distance in the volume element $|\mathbf{x}_{ba}|$ (which we call the "coupling radius"). The ejecta, in order to reach that point, must sweep up the mass m_b (in e.g. a shock or shell) – it cannot simply "spread uniformly" throughout the volume. This means some *PdV* work must have been done, converting thermal energy into kinetic energy. Thus the correct momentum ($\Delta \mathbf{p}_b''$) to couple into the domain b is *not* the initial ejecta momentum $\Delta \mathbf{p}_b'$. Rather, if the shock is energy-conserving, it is trivial to show that the correct momentum is $\Delta \mathbf{p}_b'\,(1 + m_b/\Delta m_b)^{1/2}$. In the early stages of SNe expansion, the shocks are indeed energy-conserving to high accuracy. Of course, at sufficiently long times (or equivalently large radii and/or large entrained masses), the shock becomes radiative, the residual thermal energy is lost, and the shock asymptotically reaches a final "terminal momentum" $p_{\rm t}$ (and becomes momentum, rather than energy conserving). Therefore we must impose an upper

[27] For our MFM hydrodynamic method, the face $\mathbf{A}_{ba}$ is defined as (see Hopkins 2015):

$$\mathbf{A}_{ba} \equiv \bar{n}_a^{-1}\,\bar{\mathbf{q}}_b(\mathbf{x}_a) + \bar{n}_b^{-1}\,\bar{\mathbf{q}}_a(\mathbf{x}_b) \tag{D5}$$

$$\bar{\mathbf{q}}_b(\mathbf{x}_a) \equiv \mathbf{E}_a^{-1}\cdot\mathbf{x}_{ba}\,W(\mathbf{x}_{ba}, H_a) \tag{D6}$$

$$\mathbf{E}_a \equiv \sum_c (\mathbf{x}_{ca}\otimes\mathbf{x}_{ca})\,W(\mathbf{x}_{ca}, H_a) \tag{D7}$$

For SPH, the face is defined by the simpler relation $\mathbf{A}_{ba} = [\bar{n}_a^{-2}\,\partial W(|\mathbf{x}|_{ba}, H_a)/\partial|\mathbf{x}|_{ba} + \bar{n}_b^{-2}\partial W(|\mathbf{x}|_{ba}, H_b)/\partial|\mathbf{x}|_{ba}]\,\hat{\mathbf{x}}_{ba}$. In moving-mesh or fixed-grid finite-volume codes, the face $\mathbf{A}_{ba}$ is the explicit geometric mesh face between cells.

limit $\Delta\mathbf{p}'_b\, p_t/p_{ej}$. We therefore have:

$$\Delta\mathbf{p}''_b \equiv \Delta\mathbf{p}'_b\,\mathrm{MIN}\left[\sqrt{1+\frac{m_b}{\Delta m_b}}\,,\,\frac{p_t}{p_{ej}}\right] \tag{D22}$$

$$\frac{p_t}{M_\odot\,\mathrm{km\,s^{-1}}} \approx 4.8\times10^5\left(\frac{E_{tot,ej}}{10^{51}\,\mathrm{erg}}\right)^{\frac{13}{14}}\left(\frac{n_b}{\mathrm{cm^{-3}}}\right)^{-\frac{1}{7}} f(Z_b)^{\frac{3}{2}} \tag{D23}$$

$$f(Z) = \begin{cases} 2 & (Z/Z_\odot < 0.01) \\ (Z/Z_\odot)^{-0.14} & (0.01 \le Z/Z_\odot) \end{cases} \tag{D24}$$

The expression for p_t comes from high-resolution simulations of individual SNe explosions (see e.g. Cioffi et al. 1988; Draine & Woods 1991; Slavin & Cox 1992; Thornton et al. 1998; Martizzi et al. 2015; Walch & Naab 2015; Kim & Ostriker 2015; Haid et al. 2016; Iffrig & Hennebelle 2015; Hu et al. 2016; Li et al. 2015), in media with different densities and metallicities. We discuss this at length in Paper II, and show that (1) it is the correct expression for single SNe explosions in homogeneous backgrounds, given the *same* cooling functions and all other physics implemented in FIRE, (2) it appears to be remarkably robust, across many numerical studies, and (3) our conclusions are robust to variations in the exact value of p_t much larger than its actual physical uncertainty. It is easy to verify, given the form of Eq. D22, that at sufficiently high resolution ($m_b \ll 1000\,M_\odot$), the p_t term simply never enters our equations – in other words, the SNe cooling radii are always resolved. This motivates our SNe explosion resolution criteria in the text. However the design of the expressions here is such that our coupling scheme automatically correctly treats each of e.g. the ejecta free-streaming, Sedov-Taylor, and snowplow phases. Note that since ΔE represents the *total* energy, this is not directly modified by changing $\Delta\mathbf{p}$ (the correct thermal-kinetic breakdown is automatic).

(vii) Add final fluxes to the neighboring gas elements, in a fully-conservative manner:

$$m_b^{new} = m_b + \Delta m'_b \tag{D25}$$

$$(Z\,m_b)^{new} = Z^{new}\,m_b^{new} = (Z\,m_b) + \Delta m'_{Z,b} \tag{D26}$$

$$\mathbf{p}_b^{new} = m_b^{new}\,\mathbf{v}_b^{new} = \mathbf{p}_b + \Delta\mathbf{p}''_b \tag{D27}$$

$$E_b^{new} = E_{kinetic}^{new} + U_{internal}^{new} = E_b + \Delta E'_b \tag{D28}$$

So (like our hydrodynamic update), we add conserved quantities (m, $\mathbf{p}$, E) and from those update primitive quantities (Z, $\mathbf{v}$, internal energy U, etc.). We check that any residual momentum or mass (from e.g. round-off error) is re-assigned to the star so conservation is always machine-accurate.

To be fully consistent with the radiative losses described above (when the cooling radius is un-resolved), we must also modify $U_{internal}^{new}$. Following Thornton et al. (1998), the thermal post-shock energy outside R_{cool} decays rapidly as $\propto (r/R_{cool})^{-6.5}$; so we estimate the effective R_{cool} via the requirement that, at the end of the energy-conserving phase, $(1/2)\,(m_{ej} + m_{swept}[R_{cool}])\,v_f^2 = (1/2)\,m_{ej}\,v_{ej}^2$ and $p_t = m_{swept}[R_{cool}]\,v_f$ (where m_{swept} is the enclose mass "swept up" by the shell), giving $R_{cool} \approx 28.4\,\mathrm{pc}\,(n_b/\mathrm{cm^{-3}})^{-3/7}\,(E_{tot,ej}/10^{51}\,\mathrm{erg})^{2/7}\,f(Z_b)$ for p_t in Eq. D23. If $|\mathbf{x}_{ba}| < R_{cool}$, we leave $U_{internal}^{new}$ un-modified. If $|\mathbf{x}_{ba}| > R_{cool}$, we calculate the increase in internal energy from shock-heating, ignoring cooling: $\Delta U_b \equiv U_{internal}^{new} - U_{internal}^{old}$, and then modify it to determine the correct internal energy: $U_{internal} = U_{internal}^{old} + \Delta U_b\,(|\mathbf{x}_{ba}|/R_{cool})^{-6.5}$. We show in Paper II that this extra step has a negligible effect, since (by definition), when we couple the ejecta to a size/mass scale larger than R_{cool}, it will radiate its energy rapidly, so in practice we find that if we simply leave ΔU_b un-modified, the residual energy is (correctly) radiated away in the next timestep. But for the sake of physical consistency and accuracy, we adopt the full expression here.

APPENDIX E: ALGORITHMIC IMPLEMENTATION OF RADIATIVE FEEDBACK

Now we describe the implementation of radiative feedback, used for radiation pressure (in all wavebands, UV-through-IR), photo-ionization, and photo-electric heating. The algorithm here was first developed in Hopkins et al. (2012c), and the version used in FIRE-1, which is almost exactly identical to that here,[28] was presented in Hopkins et al. (2014). As noted in the text, we for convenience denote the radiative transport algorithm as the "LEBRON" (Locally Extincted Background Radiation in Optically-thin Networks) approximation.

We emphasize that this is *not* the same as the algorithm used in some earlier work (e.g. Hopkins et al. 2011). That algorithm was developed for very specific simulations which followed *only* infrared multiple-scattering radiation pressure (ignoring single-scattering radiation pressure, photo-heating, SNe, and OB/AGB winds), with much higher resolution than the FIRE simulations here (following star formation down to protostellar cores with densities $\sim 10^6\,\mathrm{cm^{-3}}$ and size scales $< 0.1\,\mathrm{pc}$).

In Paper III, we discuss each piece of the FIRE radiative feedback algorithm in detail, and consider a large suite of both idealized test problems and full cosmological simulations, to test and validate each and show exactly how it does or does not influence our predictions. We also compare to a limited set of "fully radiation hydrodynamics" simulations (using the M1 moments method to explicitly integrate the radiative transfer equations through the code), and show that this makes no significant difference to our predictions compared to the computationally more convenient approximation we describe here.

The complete algorithm is:

(i) Determine Background Radiation (Source Luminosities): every timestep Δt_a for each star particle a, we take the luminosity $L_\nu^a = \Psi_\nu^a\,m_{*,a}$ as a function of the star particle age, metallicity, and mass, directly from the stellar evolution models. This follows the tabulation given in Appendix A, for each of five broad bands we follow: ionizing (L_{ion}, $\lambda < 912\,$Å), far-UV (L_{FUV}, $912\,\text{Å} < \lambda < 1550\,$Å), near-UV ($L_{UV}$, $1550 < \lambda < 3600\,$Å), optical/near-IR (L_{Opt}, $3600\,\text{Å} < \lambda < 3\,\mu$), and mid/far-IR ($L_{IR}$, $\lambda > 3\,\mu$). The FUV band is used for photo-electric heating, while NUV/optical bands dominate the single-scattering radiation pressure (bolometric luminosity), and the IR band is reserved for light re-radiated by dust.

(ii) Locally Extinct: We now process the absorption/extinction in the vicinity of each source. Along a sightline, the optical depth seen by the source is $\tau_\nu^a = \kappa_\nu\,\Sigma_{column}^a$, where κ_ν is the flux-weighted opacity (given in Appendix A; these are calculated for the mean un-obscured spectrum in the stellar populations models, integrated over each band). Since we are interested in *local* extinction, we approximate Σ_{column}^a using the Sobolev approximation for

[28] The only difference between the radiation algorithms in FIRE-1 and FIRE-2, as noted in the main text, is that in FIRE-2 we allow ionizing photons to propagate outside of the numerical domain boundaries. In FIRE-1, for numerical convenience, their propagation was "truncated" at these (large-scale) boundaries. The effects of this are negligible in our simulations here, because almost all the "work" done by ionizing photons is on nearby gas.

the isotropic (angle-averaged) column density integrated outward from the source:

$$\tau_\nu^a = \langle\kappa\rangle_\nu \times \langle\Sigma_{\rm column}^{a,\,\rm Sobolev}\rangle_{\phi,\,\theta} \quad \text{(E1)}$$

$$\langle\kappa\rangle_\nu \equiv \frac{\int_{\rm band}\kappa_\nu\,\langle L_\nu\rangle^{\rm unabsorbed}\,d\nu}{\int_{\rm band}\langle L_\nu\rangle^{\rm unabsorbed}\,d\nu} \quad \text{(E2)}$$

$$\langle\Sigma_{\rm column}^{a,\,\rm Sobolev}\rangle_{\phi,\,\theta} \equiv \rho_a\left[h_a + \frac{\rho_a}{|\nabla\rho_a|}\right] \quad \text{(E3)}$$

Here ρ_a, $\nabla\rho_a$, and h_a are the gas density, density gradient and inter-element spacing, evaluated at location $\mathbf{x}_a$ (with the same algorithm as our usual hydrodynamics). The $\rho_a/|\nabla\rho_a|$ term accounts for the gas column integrated to infinity – it is exact for e.g. a density distribution which declines exponentially with distance from the source – while the $\rho_a h_a$ term is just the column through the local cell. Since this is isotropic, the absorbed luminosity (in a narrow band) is just $L_{\rm abs,\,\nu}^a = (1-\exp(-\tau_\nu^a))\,L_\nu^a$ and the surviving un-absorbed luminosity is:

$$L_{\rm emergent,\,\nu}^a = \exp\left(-\tau_\nu^a\right)L_\nu^a \quad \text{(E4)}$$

Recall our spectral templates include negligible "initial" luminosity in the mid/far IR band: but we assume the luminosity absorbed by *dust* – that from the FUV, NUV, and optical/near-IR bands – is immediately re-radiated in the mid/far-IR band, giving an emergent IR luminosity:

$$L_{\rm IR}^a = \sum_{\nu={\rm FUV,\,UV,\,Opt}} L_{\rm abs,\,\nu}^a = \sum_{\nu={\rm FUV,\,UV,\,Opt}} (1-\exp(-\tau_\nu^a))\,L_\nu^a \quad \text{(E5)}$$

For the ionizing band, the opacity comes from neutral hydrogen, and we must jointly solve for the ionizing state and photon absorption; we therefore treat this separately using a simple Stromgren approximation. From Step (i) above (the stellar evolution models) we have $\dot{N}_a^{\rm ion}$, the rate of production of ionizing photons ($\propto L_{\rm ion}$). Now, we take all gas elements b in the vicinity of a and sort them by increasing distance $|\mathbf{x}_{ba}|$. Beginning with the closest, we test whether it is already ionized (either because $T_b > 10^4\,$K, or because it is already tagged as a member of another HII region), and if so we move on to the next-closest particle. If it is not ionized, we calculate the ionization rate needed to fully ionize it as: $\Delta\dot{N}_b = N(H)_b\,\beta\,n_{e,b}$ (where $N(H)_b = X_H\,m_b/\mu\,m_p$ is the number of H atoms in b, $\beta \approx 3\times10^{-13}\,{\rm cm^3\,s^{-1}}$ is the recombination coefficient, and $n_{e,b}$ is the electron density assuming full ionization). If $\Delta\dot{N}_b \leq \dot{N}_a^{\rm ion}$, then particle b is tagged as being within an HII region, and the photons are "consumed," so $\dot{N}_a^{\rm ion} \rightarrow \dot{N}_a^{\rm ion} - \Delta\dot{N}_b$. We then proceed to the next particle and repeat. If we reach a particle which is not ionized but for which $\Delta\dot{N}_b > \dot{N}_a^{\rm ion}$, we determine whether or not to ionize it randomly, with probability $p = \dot{N}_a^{\rm ion}/\Delta\dot{N}_b$, and consume the remaining photons (guaranteeing the correct total mass is ionized, on average). Any particle tagged as "within an HII region" is fully ionized and not allowed to cool to temperatures lower than $< 10^4\,$K within that same timestep. If we reach the end of the local computational domain, or distance from the source where the optically-thin flux density falls below the meta-galactic ionizing background, we stop the iteration and the remaining photons are "emergent" as in Eq. E4.

(iii) Account for Momentum of Locally-Absorbed Photons: Over a timestep Δt, the absorbed photons in these bands impart their single-scattering momentum to the surrounding gas, radially directed away from the star particle, with total momentum

$$\Delta p = \frac{L_{\rm abs}^a}{c}\,\Delta t = \frac{\Delta t}{c}\sum_\nu L_{\rm abs,\,\nu}^a \quad \text{(E6)}$$

This momentum flux is distributed among the neighbors, directed radially away from the star particle, as described in Appendix D.

(iv) Transport the Locally-Extincted Radiation via an Optically-thin Network: We now have an "emergent" spectrum *after local attenuation* around each star, $L_{\rm emergent,\,\nu}$. Since we assume the absorption is dominated by the gas/dust local to the star, and the emission (from the star) is isotropic, the incident flux $\mathbf{F}_\nu^b$ and photon energy density e_ν^b at a distant gas element b are just

$$\mathbf{F}_\nu^b = \sum_a \mathbf{F}_{\nu,\,\rm emergent}^a = \sum_a \frac{L_{\rm emergent,\,\nu}^a}{4\pi\,|\mathbf{x}_b-\mathbf{x}_a|^2}\,\frac{\mathbf{x}_b-\mathbf{x}_a}{|\mathbf{x}_b-\mathbf{x}_a|} \quad \text{(E7)}$$

$$e_\nu^b = \sum_a \frac{L_{\rm emergent,\,\nu}^a}{4\pi\,c\,|\mathbf{x}_b-\mathbf{x}_a|^2} \quad \text{(E8)}$$

This is identical in form to the equation for gravity, so is computed in the same pass in the gravity tree (we "soften" the sources identical to how we soften gravity, in fact, to prevent a $1/r^2$ divergence and reflect the physical fact that each star particle really represents many stars distributed within the softening length).

(v) Calculate incident radiative acceleration from long-range fluxes: For a gas element b with effective face area A_b and mass m_b (hence surface density $\Sigma_b \equiv m_b/A_b$) seeing an incident flux $\mathbf{F}_\nu^b$, the exact radiative acceleration is given by

$$\left.\frac{d\mathbf{v}_b}{dt}\right|_\nu = \frac{1}{m_b}\left.\frac{d\mathbf{p}_b}{dt}\right|_\nu = \frac{\mathbf{F}_\nu^b}{c}\,\frac{A_b}{m_b}\left[1-\exp\left(-\kappa_\nu\,\frac{m_b}{A_b}\right)\right] \quad \text{(E9)}$$

In the optically thin limit ($\kappa_\nu\,\Sigma_b \ll 1$), this reduces to the common expression $\mathbf{a} = \kappa_\nu\,\mathbf{F}_\nu/c$, but in the optically thick limit ($\kappa_\nu\,\Sigma_b \gg 1$), the force saturates at $m_b\,\mathbf{a} = (\mathbf{F}_\nu\,A_b)/c$, i.e. the element absorbs all the flux across its subtended area (but no more). For simplicity here we take the effective area to be that of a sphere with the same volume as the element ($= m_b/\rho_b = (4\pi/3)\,h_b^3$), i.e. $\pi\,h_b^2$; using the more complicated hydrodynamic face areas introduces negligible ($\ll 10\%$) differences in the accelerations here. We adopt $\kappa_\nu = \langle\kappa\rangle_\nu$ for each band.

(vi) Self-shield and pass incident fluxes to cooling routines: Because we have accounted for shielding around the emitter, but not the absorber, we include an additional shielding pass for the photo-heating terms: at a gas element b, we take the photon energy density $\langle e_\nu^b\rangle = e_\nu^b\,\exp(-\tau_\nu^b)$, where e_ν^b is the photon energy density given by Eq. E8, and τ_ν^b is the optical depth estimated using the Sobolev approximation in Eq. E1, but now at the location of the absorbing *gas* element (instead of around the emitting star). The resulting, shielded e_ν^b are passed to the cooling/heating routines, to compute photo-ionization and photo-electric heating rates as in Appendix B.

APPENDIX F: ADDITIONAL PHYSICS: MAGNETIC FIELDS, CONDUCTION, VISCOSITY, TURBULENT DIFFUSION, COSMIC RAYS, RADIATION HYDRODYNAMICS, AND BLACK HOLES

Here, we describe the numerical implementations of additional physics *not* included in the "core physics only" FIRE simulations, but studied in companion papers (e.g. Su et al. 2016) which take standard FIRE-2 simulations and add, e.g. magnetic fields. We emphasize again that these physics are not used in the "default" or "core physics" runs in this paper. However, because we wish to present a complete, thorough, and fully-consistent set of numerical methods, we summarize them here, referring to the appropriate methods papers for more details.

F1 Magnetic Fields

In simulations with magnetic fields, we solve the equations of ideal magnetohydrodynamics (MHD) as implemented in GIZMO in Hopkins & Raives (2016). The exact numerical formulation of the equations is presented there along with an extensive series of several dozen test problems, as well as tests of full galaxy simulations using our FIRE physics. The tests demonstrate that the implementation in our MFM solver is accurate and converges at second order. In particular, Hopkins & Raives (2016), Hopkins (2016a), and Hopkins & Lee (2016) show that our implementation correctly captures traditionally difficult phenomena such as the magneto-rotational instability (MRI), magnetic jet launching in disks, magnetic fluid-mixing instabilities, and sub-sonic and super-sonic MHD turbulent dynamos. The accuracy and convergence order appears comparable to state-of-the-art grid codes (e.g. ATHENA) on the problems of interest and greatly superior to the P-SPH implementation in GIZMO, especially in problems where angular momentum, super-sonic advection, strong shocks, and fluid mixing instabilities appear (typical of cosmological simulations). Non-ideal MHD effects (Ohmic resistivity, ambipolar diffusion, and the Hall effect) are also implemented and well-tested in GIZMO (Hopkins 2016b), but these are not expected to be important on galactic scales.

Readers interested in more details of our numerical implementation of MHD should consult Hopkins & Raives (2016) and the public GIZMO source code.

F2 Anisotropic Spitzer-Braginskii Viscosity and Conduction

The implementation of anisotropic Spitzer-Braginskii viscosity and conduction in GIZMO is described and tested in Hopkins (2016b). In addition to the usual MHD fluxes, this adds an anisotropic viscous stress-energy tensor $\mathbf{\Pi}$ to the momentum flux ($\mathbf{F_p} = \mathbf{\Pi}$) and energy flux ($\mathbf{F}_e = \mathbf{\Pi}\cdot\mathbf{v}$), and a conductive energy flux $\mathbf{F}_e = \mathbf{K}\cdot\nabla T$. The appropriate anisotropic tensor expressions for MHD are given by (Spitzer & Härm 1953; Braginskii 1965):

$$\mathbf{K} \equiv \kappa_{\rm cond}\,\hat{B}\otimes\hat{B} \tag{F1}$$

$$\kappa_{\rm cond} = \frac{0.96\,f_i\,(k_B T)^{5/2}\,k_B}{m_e^{1/2}\,e^4\,\ln\Lambda}\,\left(1+4.2\,\ell_e/\ell_T\right)^{-1} \tag{F2}$$

$$\mathbf{\Pi} \equiv 3\,\nu_{\rm visc}\left(\hat{B}\otimes\hat{B}-\frac{1}{3}\mathbf{I}\right)\left[\left(\hat{B}\otimes\hat{B}-\frac{1}{3}\mathbf{I}\right):(\nabla\otimes\mathbf{v})\right] \tag{F3}$$

$$\nu_{\rm visc} = \frac{0.406\,f_i\,m_i^{1/2}\,(k_B T)^{5/2}}{(Z_i\,e)^4\,\ln\Lambda}\,\left(1+4.2\,\ell_i/\ell_v\right)^{-1} \tag{F4}$$

where $\otimes$ denotes the outer product; $\hat{B}$ is the direction of the magnetic field vector; $\mathbf{I}$ is the identity matrix; $\mathbf{v}$ the velocity; ":" denotes the double-dot-product ($\mathbf{A}:\mathbf{B} \equiv {\rm Trace}(\mathbf{A}\cdot\mathbf{B})$); $\ln\Lambda \approx 37.8$ is the Coulomb logarithm (Sarazin 1988); m_e, e, m_i, $Z_i\,e$ are the electron mass and charge and ion mass and charge; f_i the ionized fraction (calculated self-consistently in our cooling routines); k_B the Boltzmann constant; ℓ_e (ℓ_i) is the electron (ion) mean-free path, and $\ell_T = T/|\nabla T|$ ($\ell_v = |\mathbf{v}|/||\nabla\otimes\mathbf{v}||$) is the temperature (velocity) gradient scale length (this term correctly accounts for saturation of $\kappa_{\rm cond}$ or $\nu_{\rm visc}$ when electrons/ions have long mean free paths, by not allowing the gradient scale length to be shorter than $\ell_{e,i}$). In these equations, $\kappa_{\rm cond}$ is the conductivity, and $\nu_{\rm visc}$ the viscosity. Additional details of the coefficients, and a study of their effects, are in Su et al. (2016). In Hopkins (2016b), we show that the numerical implementation of these fluxes is accurate, able to handle arbitrarily large anisotropies, converges comparably to higher-order fixed-grid codes, and is able to correctly capture complicated non-linear instabilities sourced by anisotropic diffusion such as the magneto-thermal and heat-flux bouyancy instabilities.

Readers interested in more details of our numerical implementation of anisotropic diffusion should consult Hopkins (2016b) and the public GIZMO source code.

F3 Sub-Grid Turbulent Eddy Diffusivity

In some models for turbulence (e.g. mixing-length theory), the effects of unresolved (small-scale) eddies and microphysical processes transporting passive scalars (such as metals) are treated as diffusive processes. The implementation and tests of these diffusion operators in GIZMO are presented in Hopkins (2016b); the solver is an explicit finite-volume scheme which converges at second-order accuracy (comparable to higher-order grid codes) and manifestly conserves metal mass. For a passive scalar, the transport equation is: $\partial(\rho Z)/\partial t = \nabla\cdot(\kappa_{\rm turb}\,\rho\,\nabla Z)$, where Z is the abundance per unit mass of the scalar (i.e. the metallicity) and the "eddy diffusivity" $\kappa_{\rm turb} \sim \lambda_{\rm eddy}\,v_{\rm eddy}$ is the product of the scale length and rms velocity of the largest un-resolved eddies (those at the resolution scale), which dominate the transport on unresolved scales (larger eddies are, of course, resolved). In other words, one simply assumes that the diffusion or mixing time at scale $\lambda_{\rm eddy}$ scales with the eddy turnover time. Following the common Smagorinsky (1963) approximation, we can approximate the "eddy diffusivity" as

$$\kappa_{\rm turb}^{a} \equiv \sqrt{2}\,C^2\,\|\mathbf{S}_a\|\,h_a^2 \tag{F5}$$

where $C \sim 0.05-0.15$ is a constant calibrated to numerical simulations, motivated by a Kolmogorov cascade in Smagorinsky (1963), h_a is the grid scale (for our MFM method, this is equal to the rms inter-element separation), and $\mathbf{S}_a \equiv [(\nabla\otimes\mathbf{v})_a + (\nabla\otimes\mathbf{v})_a^T] - {\rm Trace}(\nabla\otimes\mathbf{v})_a/3$ is the symmetric shear tensor (and $\|\mathbf{S}\|$ denotes the Frobenius norm). Note that we use our higher-order matrix-based gradient formalism from Hopkins (2015) to calculate $\mathbf{S}_a$; this is much more accurate and less noisy compared to common SPH or pure "face-based" mesh gradient estimators, which is important to reduce artificial numerical diffusivity (see e.g. Maron & Howes 2003; Luo et al. 2008; Lanson & Vila 2008; Mocz et al. 2014; Pakmor et al. 2016).

The key assumption here – namely, the assumption that the diffusion timescale scales with eddy turnover time – has been verified in many experiments on ISM scales (see Pan & Scannapieco 2010; Petit et al. 2015, and references therein), and the scaling in Eq. F5 has been used in many applications in galaxy simulations (e.g. Wadsley et al. 2008). In Colbrook et al. (2016), we have performed our own study of the turbulent mixing, using 3D, high-resolution supersonic turbulent box simulations (with and without magnetic fields and/or shear), and verify that this prescription, with $C \approx 0.05$, is reasonable specifically in our identical MFM code with the definitions of h and $\mathbf{S}$ here (although such simple prescriptions do fail to capture some potentially important non-Gaussian features which emerge from real, resolved turbulent mixing). An independent, more extensive study (including a range of more complex problem setups) will be presented in Rennehan et al. (in prep.), but also finds $C \approx 0.03-0.05$. We therefore adopt $C = 0.05$. However, in the main text (§ 7.2) and in Escala et al. (in prep.), we show that order-of-magnitude variation in C produces no significant effects on our predictions.

We stress that a term like Eq. F5 is "built into" many numerical hydrodynamic methods. Specifically, it is well known that in finite-volume methods with advective mass fluxes (e.g. traditional grid-based methods or moving-mesh codes), an intrinsic numerical diffusivity in advection with magnitude $\sim h_a\,\Delta v(h_a)$ appears; this

automatically produces scalar/metal diffusion via "numerical mixing." It is straightforward to show that if the necessary assumptions of the Smagorinsky model (Eq. F5) are true, then the artificial numerical mixing in these methods is always *larger* than the "true" effective turbulent diffusivity. If we used such a method, it would therefore not be necessary to explicitly solve Eq. F5. However, in our default MFM hydrodynamic method, we follow fixed-mass elements (i.e. there are no advective mass fluxes, hence no artificial "numerical diffusivity" of passive scalars). While the methods will converge to an identical solution at sufficiently high resolution (Hopkins 2015), the concern is that at fixed resolution, MFM will under-estimate the metal-mixing owing to un-resolved small-scale eddies that should mix between the boundaries of neighboring resolution elements. If we resolved individual stars, the stars would draw mass from many resolution elements (each with their own abundances) and this would still not be a problem, but at our resolution single star particles inherit the abundances of their (single) parent gas particle, so this effect can artificially introduce "shot noise" in the abundances of stars forming from neighboring gas elements if we do not include an explicit numerical mixing term. One therefore can view Eq. F5 as a purely numerical term which "restores" the desirable aspect of the numerical diffusivity present in certain numerical methods.

We caution, however, that simple diffusion prescriptions such as Eq. F5, naively applied, can substantially over-estimate the diffusivity. The critical assumption is that the resolution-scale motion $\Delta v \sim \|\mathbf{S}\|\,h \sim v_{\rm eddy}$ is entirely due to turbulence; if there is any bulk motion included in $\mathbf{S}$, this will over-estimate $\kappa_{\rm turb}$. This can be particularly problematic if e.g. differential rotation in a disk or shear in CGM outflows is poorly-resolved, in which case the naively inferred $\kappa_{\rm turb}$ can over-estimate by an order of magnitude the true turbulent motion (for example, if the disk scale height $H_{\rm disk} \approx \sigma_{\rm turb}/\Omega$ is unresolved, $h_a \gg H_{\rm disk}$, then Eq. F5 will return $\kappa_{\rm turb} \sim h_a\,(h_a\,\Omega)$ instead of the correct maximum diffusivity for disk-scale eddies, $\kappa_{\rm turb} \sim H_{\rm disk}\,\sigma_{\rm turb} \sim H_{\rm disk}^2\,\Omega$). There is no obvious universal "switch" to cure these pathologies; however we can limit the magnitude of the errors. In Hopkins (2015) we develop two mesh-free finite-element hydrodynamic methods, our default MFM method here, and a "meshless finite-volume" (MFV) method, the latter of which includes advective mass fluxes (more similar to a moving-mesh code), but otherwise is identical to MFM. The MFV method therefore includes the inherent numerical diffusivity described above. In the diffusion step, therefore, we can first calculate the absolute value of the metal flux that would have been calculated in MFV (owing simply to the advection term; see Hopkins 2015 for the exact values of these terms), and then impose this as an upper limit to the diffusive flux. Since MFV is a second-order, quasi-Lagrangian method, this eliminates the most egregious errors in Eq. F5. We find that in idealized test problems, this correction is negligible, but in realistic cosmological simulations it prevents the most severe pathological situations. Essentially, then, our implementation of "unresolved turbulent diffusion" (Eq. F5) is guaranteed to – *at most* – produce the same metal mixing we would have obtained had we simply run our simulations using a finite-volume (MFV or moving-mesh) hydrodynamic method. A more detailed study of various (more sophisticated) turbulent and numerical mixing models will be presented in Rennehan et al. (in prep.); preliminary results indicate that the alternative methods give similar results in galaxy-scale simulations.

Readers interested in more details of our numerical implementation of eddy diffusivity should consult Hopkins (2016b) and the public GIZMO source code.

F4 Cosmic Rays

A detailed series of papers on our numerical treatment of cosmic rays (CRs) and their consequences for galaxy and star formation will be the subject of future work. Here we briefly outline the key physical equations being solved.

CRs are approximated as a single-spectrum, ultra-relativistic ($\gamma_{\rm cr} = 4/3$) fluid. By "single-spectrum," we refer to an invariant *shape* of the CR energy distribution (only its normalization varies); the equations of motion are then integrated over this spectrum. Because low-energy CRs dominate the energy density and pressure (hence effects on the gas), this is essentially equivalent to a single-species treatment of the CRs with energy $\sim$GeV.

In the fluid limit, for a CR energy density $e_{\rm cr}$ the corresponding pressure is $P_{\rm cr} = (\gamma_{\rm cr}-1)\,e_{\rm cr}$ and the evolution equation (temporarily neglecting sources, cooling, and diffusion) is $\partial e_{\rm cr}/\partial t = -\nabla\cdot(\mathbf{v}_{\rm cr}\,e_{\rm cr}) - P_{\rm cr}\,\nabla\cdot\mathbf{v}_{\rm cr}$ where $\mathbf{v}_{\rm cr} \equiv \mathbf{v}_{\rm gas} + \mathbf{v}_{\rm streaming}$. Here $\mathbf{v}_{\rm cr}$ is the CR mean bulk velocity, which is decomposed by standard convention into the gas velocity $\mathbf{v}_{\rm gas}$ (simple advection with gas) and the streaming velocity $\mathbf{v}_{\rm stream}$ (bulk motion relative to the gas). The first term $\nabla\cdot(\mathbf{v}_{\rm cr}\,e_{\rm cr})$ is just the continuity equation; the second $P_{\rm cr}\,\nabla\cdot\mathbf{v}_{\rm cr}$ represents adiabatic "PdV" work. Following Uhlig et al. (2012), in a Lagrangian method such as ours, it is convenient to re-write this in terms of the Lagrangian derivative co-moving with the *gas* (since we solve the CR fluid equations across gas elements), $d/dt = \partial/\partial t + \mathbf{v}_{\rm gas}\cdot\nabla$. With that definition and some re-arrangement, we obtain:

$$\begin{aligned}\rho\,\frac{du_{\rm cr}}{dt} &= -P_{\rm cr}\nabla\cdot\mathbf{v}_{\rm gas} \\ &\quad -\nabla\cdot\left[\mathbf{v}_{\rm streaming}\,(e_{\rm cr}+P_{\rm cr})\right] + \mathbf{v}_{\rm streaming}\cdot\nabla P_{\rm cr} \\ &\quad -\nabla\cdot\left[\left(\kappa_{\rm cr}\,\hat{B}\otimes\hat{B}\right)\cdot\nabla e_{\rm cr}\right] \\ &\quad -\Gamma_{\rm cr} + \rho\,\dot{u}_{\rm cr,\,*}\end{aligned} \tag{F6}$$

where $u_{\rm cr} \equiv e_{\rm cr}/\rho$ is the specific CR energy. These are solved in the Riemann problem in the same manner as the standard hydrodynamic equations; the equations are implemented identically so the accuracy and convergence rate is also identical. Just like for the hydrodynamic equations, we evolve conservative variables (mass, momentum, energy), which for the CRs is the total CR energy associated with particle a ($E_{{\rm cr},a} = u_{{\rm cr},a}\,m_a$), such that our method is manifestly conservative.

The first term $-P_{\rm cr}\nabla\cdot\mathbf{v}_{\rm gas}$ represents adiabatic compression of the CR fluid, and is handled in the standard hydrodynamic manner. The second and third terms represent streaming transport and losses, respectively: the streaming velocity is given by:

$$\mathbf{v}_{\rm streaming} = -\left(c_s^2 + v_A^2\right)^{1/2}\,\left(\hat{B}\otimes\hat{B}\right)\cdot\frac{\nabla e_{\rm cr}}{|\nabla e_{\rm cr}|} \tag{F7}$$

The direction of $\mathbf{v}_{\rm streaming} \propto -(\hat{B}\otimes\hat{B})\cdot\nabla e_{\rm cr}/|\nabla e_{\rm cr}|$ trivially reflects the physical requirement that CRs stream down the CR number density/energy/pressure gradient, along the magnetic field lines (given by the projection operator $\hat{B}\otimes\hat{B}$). The streaming speed $v_{\rm st} = (c_s^2+v_A^2)^{1/2}$ interpolates between the known limits in the strongly-magnetized case (plasma $\beta = P_{\rm thermal}/P_{\rm magnetic} \ll 1$, $v_A \gg c_s$), where $v_{\rm st} \rightarrow v_A$, and the weak-field case ($\beta \gg 1$), where $v_{\rm st} \rightarrow c_s$. For more discussion, see Holman et al. (1979); Achterberg (1981); Felice & Kulsrud (2001); Yan & Lazarian (2008); Enßlin et al. (2011). This means the transport term $-\nabla\cdot[\mathbf{v}_{\rm streaming}\,(e_{\rm cr}+P_{\rm cr})]$ has the mathematical form of an anisotropic scalar diffusion equation: $\nabla\cdot[(\kappa_{\rm st}\,\hat{B}\otimes\hat{B})\cdot\nabla e_{\rm cr}]$ with "diffusivity" $\kappa_{\rm st} = v_{\rm st}\,(e_{\rm cr}+P_{\rm cr})/|\nabla e_{\rm cr}|$; the numerical implementation in GIZMO for these equations and extensive tests of their accuracy are presented in Hopkins (2016b).

The term $\mathbf{v}_{\rm streaming}\cdot\nabla P_{\rm cr}$ represents work done by the CRs in streaming; expanding it we obtain $=-v_{\rm st}\,|\hat{B}\cdot\nabla P_{\rm cr}|^2/|\nabla P_{\rm cr}|$, which is negative-definite (energy is always lost from the CRs). Physically, the energy goes into self-excited Alfven waves with extremely high frequencies (wavelengths of order the gyro radius, $\sim 10^{11}\,$cm at $\sim$GeV); as a result these waves damp and thermalize their energy almost instantly (Kulsrud & Pearce 1969).

The term $\nabla\cdot[(\kappa_{\rm cr}\hat{B}\otimes\hat{B})\cdot\nabla e_{\rm cr}]$ represents anisotropic diffusion of the CRs, from non-streaming processes (e.g. microturbulence). Once again the projection operator $\hat{B}\otimes\hat{B}$ appears, because diffusion is confined along field lines. The same numerical treatment of anisotropic diffusion from Hopkins (2016b) is used. The diffusivity $\kappa_{\rm cr}$ is perhaps the most uncertain element of the physics: observations suggest $\kappa_{\rm cr}\sim 10^{28}-10^{29}\,{\rm cm^2\,s^{-1}}$ in the MW, but this is highly model-dependent (see Gaggero et al. 2015, and references therein). The "textbook" scaling for $\kappa_{\rm cr}$ assumes a Kolmogorov spectrum of magnetic field fluctuations, giving (Berezinskii et al. 1990)

$$\kappa_{\rm cr}\approx\frac{c\,r_{\rm gyro}}{3}\,\frac{B^2_{\rm coherent}}{B^2_{\rm random}(r_{\rm gyro})} \qquad \text{(F8)}$$
$$\approx 3\times10^{28}\,\frac{\rm cm^2}{\rm s}\,R_{\rm GV}^{1/3}\left(\frac{\mu G}{|\mathbf{B}|}\right)^{\frac{1}{3}}\left(\frac{L_{\rm drive}}{\rm kpc}\right)^{\frac{2}{3}}$$

where $B_{\rm coherent}$ and $B_{\rm random}(r_{\rm gyro})$ represent the coherent mean $\mathbf{B}$-field and random component on the scale $r_{\rm gyro}$ of the gyro/Larmor radius of CRs, $R_{\rm GV}\approx 1$ is the magnetic rigidity in gigavolts (which determines $r_{\rm gyro}$), and $B_{\rm random}(r_{\rm gyro})$ is scaled from the resolved mean-field $|\mathbf{B}|$ assuming a Kolmogorov spectrum with driving scale $L_{\rm drive}$ (which we can approximate via the local pressure gradient scale-length $=P_{\rm tot}/|\nabla P_{\rm tot}|$).

The term $\Gamma_{\rm cr}$ represents the loss rate of CRs to collisions and radiation. We take:

$$\Gamma_{\rm cr}=7.51\times10^{-16}\,{\rm s^{-1}}\,e_{\rm cr}\,(1+0.22\,\tilde{n}_e)\left(\frac{n_{\rm H}}{\rm cm^{-3}}\right) \qquad \text{(F9)}$$

from Guo & Oh 2008, including combined hadronic plus Coulomb losses; $n_{\rm H}$ is the hydrogen number density and $\tilde{n}_e$ is the number of free electrons per hydrogen nucleus (the term in $\tilde{n}_e$ is the Coulomb term). This is implemented for the CRs identically to our normal gas-cooling operations.

The term $\dot{u}_{\rm cr,\,*}$ represents CR injection by sources. In our simulations, we model injection by strong shocks owing to SNe and fast OB-star winds. Whenever feedback energy from SNe or fast OB-star wind is coupled to gas, we assume a fraction $\epsilon_{\rm cr}<1$ of the initial ejecta kinetic energy is converted into CR energy (removed from the kinetic energy to be fully conservative), and distributed to the surrounding gas elements around the star in the identical manner to the other ejecta products (mass, metals, etc). For convenience we define a "fast" wind as one with launch velocity $>500\,{\rm km\,s^{-1}}$, but varying this has little effect on our results since the injection is completely dominated by SNe.

Finally, we need to include the CR effects on the *gas* dynamical equations. In addition to the normal MHD terms, the following CR terms appear in the gas momentum and energy equations:

$$\rho\,\frac{d\mathbf{v}_{\rm gas}}{dt}\Big|_{\rm cr}=-\nabla P_{\rm cr} \qquad \text{(F10)}$$

$$\rho\,\frac{du_{\rm gas}}{dt}\Big|_{\rm cr}=-\mathbf{v}_{\rm streaming}\cdot\nabla P_{\rm cr}+\tilde{\Gamma}_{\rm cr}^{\rm thermal} \qquad \text{(F11)}$$

$$\tilde{\Gamma}_{\rm cr}^{\rm thermal}=7.51\times10^{-16}\,{\rm s^{-1}}\,e_{\rm cr}\left(\frac{1}{6}+0.22\,\tilde{n}_e\right)\left(\frac{n_{\rm H}}{\rm cm^{-3}}\right) \qquad \text{(F12)}$$

The $\nabla P_{\rm cr}$ term just represents CR pressure accelerating the gas, i.e. the CR pressure simply adds linearly to the gas thermal pressure to give the total pressure that appears in the momentum equation; this is included directly in the momentum equation in the MHD Riemann problem (trivially implemented in the HLLD Riemann solver used in GIZMO). The $\mathbf{v}_{\rm streaming}\cdot\nabla P_{\rm cr}$ term represents the thermalized energy from the high-frequency streaming-excited waves, described above. And $\tilde{\Gamma}_{\rm cr}^{\rm thermal}$ represents the portion of the CR losses which are thermalized: following Guo & Oh (2008), all Coulomb losses and $\approx 1/6$ of the hadronic losses thermalize, with the remaining $\approx 5/6$ of the hadronic losses going into optically-thin radiation that can escape. This is included as a gas heating term self-consistently in our full heating/cooling routines.

Readers interested in details of the numerical methods for anisotropic diffusion and streaming should consult Hopkins (2016b).

F5 Radiation Hydrodynamics

All of our simulations include some approximate treatment of radiative transfer. In Appendix E, we describe our default FIRE model, and in Paper III we compare it to an alternative model. However, we note that GIZMO now includes a number of different radiation hydrodynamics (RHD) solvers. This includes flux-limited diffusion (FLD), the optically-thin variable Eddington tensor method (OTVET; Gnedin & Abel 2001), the M1 moment solver (M1; Levermore 1984), explicit ray-tracing, and the method from Hopkins et al. (2012c) which we refer to as the "Locally-Extincted Background Radiation in Optically-thin Networks" (LEBRON) method. We can, in principle, replace our default treatment of the radiation with any of these methods.

We emphasize that the choice of solver only affects the transport step for radiation. The choice of which bands/frequencies are explicitly evolved (Appendix E), the corresponding opacities and source functions from stars for each of these bands (Appendix A), and the physics of the radiation-gas coupling (the radiation pressure, photo-ionization and photo-electric heating terms; Appendix B & E) are independent of the transport method (e.g. LEBRON or M1 or ray-tracing).

Readers interested in more details of our RHD methods should see Paper III; numerical tests of the anisotropic radiation diffusion and streaming algorithms are in Hopkins (2016b), and a detailed comparison of these methods on various RHD test problems will be presented in future work.

F6 Black Holes

In future work we will explore the role of super-massive black holes (SMBHs) in galaxy evolution. The detailed physics of SMBH seed formation, subsequent growth and evolution, and feedback remain highly uncertain (both theoretically and observationally), so the future work will necessarily explore different physical models. Here we simply outline the basic algorithmic elements. Examples of the black hole algorithms currently implemented in GIZMO (but applied to non-FIRE simulations) can be found in Anglés-Alcázar et al. (2017).

In GIZMO, each SMBH is tracked via an individual, collisionless sink particle as in Springel et al. (2005). These SMBH sink particles are spawned in a manner similar to star particles, according to some rule, e.g. a probability $p_{\rm seed}=p(...)$ which depends on some set of properties of individual gas or star particles (of course, pre-existing SMBH particles can be directly inserted into the initial conditions). The gravitational softening of the SMBH particle is set to some small, fixed physical value $\epsilon_{\rm BH}$, to limit hard-scattering; we

typically chose this value to correspond to radii where the structure around the SMBH can no longer be resolved.

Once formed, each SMBH particle a has three associated mass elements. (1) The total dynamical (particle) mass m_a, which represents all mass associated with the BH and bound to it (e.g. unresolved stars and/or gas as well). This is the mass which enters the gravity equations. (2) The "accretion disk mass" $m_{\rm disk,\,a}$, which represents gas specifically accreted into the disk, but not yet into the BH. (3) The black hole mass $m_{\rm BH,\,a}$, which is the mass of the actual BH. For some purposes (e.g. gravity) this sub-division of the mass is irrelevant; for others (e.g. modeling BH spin evolution) it is necessary.

Average properties of gas and stars surrounding the SMBH particle are computed via mass-weighted kernel averages, e.g. $\langle f \rangle_a \equiv (\sum_b f_b\,m_b)/(\sum_b m_b)$ where the sum is over neighbors b within some distance $|\mathbf{x}|_{ba} < H_a$, where H_a is set as in the hydrodynamic operator to include some effective neighbor number $N_{\rm ngb,\,BH}$ (typically set to $N_{\rm ngb,\,BH} \approx 256$, much larger than the hydrodynamic equivalent, to enable computation of more complex quantities such as the angular momentum of surrounding particles). To prevent spurious "overreach" we enforce an upper limit $H_a < H_{\rm max,\,BH}$.

To ensure exact mass conservation in the simulations, and fixed particle masses, the total SMBH particle mass m_a grows only via explicit "capture" or "swallowing" of individual elements m_b (which are then deleted from the simulation). All conserved quantities are transferred to the SMBH particle. Gas elements are captured according to some rule $p_{\rm swallow}(...)$. Collisionless (DM, star, or other SMBH) particles b are captured if they (1) fall within the interaction kernel $|\mathbf{x}|_{ba} < H_a$, (2) are gravitationally bound to the SMBH particle, $|\mathbf{v}_b - \mathbf{v}_a|^2 < 2\,G\,(m_a+m_b)/|\mathbf{x}|_{ba}$, and (3) have apocentric radii (calculated assuming a two-body orbit between a and b with their instantaneous relative velocity) $r_{\rm apo} < 10\,\epsilon_{\rm BH}$. The accretion disk reservoir grows via accretion according to some function $\dot{m}_{\rm disk,\,a} = \dot{m}^{\rm acc}_{\rm disk}(m_a, ...)$ which may depend on both the SMBH particle and neighbor properties around the SMBH. The accretion disk loses mass to three sources. (1) Accretion onto the SMBH according to some function $\dot{m}^{\rm BH}_{\rm disk,\,a} = -(1-\epsilon_r)\,\dot{m}_{\rm BH-disk}(m_{\rm disk}, ...)$ which generally depends only on the accretion disk properties. This mass $\Delta m_{\rm BH,\,a} \equiv |\dot{m}^{\rm BH}_{\rm disk,\,a}|\,\Delta t_a$ is added to the BH mass reservoir $m_{\rm BH,\,a}$. (2) Radiative losses $\dot{m}^{\rm rad}_{\rm disk,\,a} = -\epsilon_r\,\dot{m}_{\rm BH-disk}(m_{\rm disk}, ...)$ (here $\epsilon_r \sim 0.1$ is the BH radiative efficiency), which are removed from the mass reservoir but determine the radiative feedback budget. (3) Accretion-disk winds $\dot{m}^{\rm wind}_{\rm disk,\,a} = -\dot{m}_{\rm disk-wind}(m_{\rm disk}, ...)$, which couple to gas around the SMBH.

Given some non-zero accretion rate, the BH can act on the gas via one of several feedback channels. Like stars, BHs exert radiative feedback in the form of radiation pressure, photo-ionization, Compton, and photo-electric heating; they can exert mechanical feedback in the form of jets or accretion-disk winds; and they can inject cosmic rays. The relevant inputs (radiative and kinetic luminosities) are given by some function of the disk properties and BH accretion rate, as specified above. By default, the feedback *coupling* is handled via the same algorithms described above for stellar feedback; however for certain applications (e.g. relativistic jets) we will explore alternative coupling algorithms.

The SMBH particle obeys all the usual time-step constraints for collisionless particles. As for star particles, we impose additional timestep restrictions to ensure the physics are time-resolved. We require $dt_a < 0.001\,m_{\rm BH,\,a}/\dot{m}_{\rm BH,\,a}$ (limiting the BH growth and integrated feedback energy deposition in a single timestep), and $dt_a < {\rm MIN}(dt_b)$ where dt_b is the timestep of all neighbor particles inside the BH interaction kernel (H_a above), so that the BH cannot "skip past" any evolution of any neighbor particle with whom it interacts.

Future studies which explore SMBH physics will include the explicit details of each algorithm above, as implemented in that study (e.g., Anglés-Alcázar et al., in prep.).

APPENDIX G: COMPUTATIONAL SCALING & RUNTIME REQUIREMENTS

The work here was made possible by extensive optimizations and improvements to the code scaling for "zoom-in" simulations. The challenge in these high-dynamic range problems is that small, dense regions (e.g. dense GMCs or star clusters), which occupy an extremely small fraction of the total mass and volume of the simulation (and therefore cannot be "broken up" over too many processors) require extremely small timesteps. But, especially with strong feedback present, the rest of the simulation (which is free to take much larger timesteps) cannot advance until they "catch up." We have addressed this with several improvements.

(i) We have made optimizations to the feedback sub-routines (for example, long-range radiation forces) which require significant neighbor communication but can be efficiently included in other operations such as tree construction, reducing their portion of the runtime from $\approx 40\%$ in previously-published work to $< 10\%$ in the current code. More generally, the entire GIZMO code has been line-by-line optimized, manually unrolling or (where possible) vectorizing certain expensive operations, replacing expensive functions in neighbor loops with look-up tables, eliminating redundant operations, and pre-computing additional quantities outside of loops; this has produced an additional factor ~ 2 speed improvement.

(ii) We have optimized the structure of the domain decomposition to make it more spatially flexible and separately parallelize each level of the timestep hierarchy (increasing the memory imbalances by factors of $\sim 2-3$, but extending the strong scaling to ~ 2 times as many cores at fixed resolution). We have also more aggressively implemented a problem-specific particle weighting scheme, where e.g. dense, star-forming gas, and young stars (as opposed to old stars) are given larger weights in the domain decomposition so that their future cost (via gravitational collapse and/or feedback) is more accurately predicted. This allows for a further factor of $\sim 2-3$ reduction in load imbalances at the smallest timesteps.

(iii) We employ a hybrid tree-particle mesh gravity solver, following GADGET-3, to efficiently reduce the cost of the gravity solution for the low-resolution regions outside of the zoom-in area.

(iv) We use the adaptive individual-timestep integration scheme from GADGET-3 with a hierarchical power-of-two subdivision, updated such that in each timestep we calculate pairwise updates of all fluxes of conserved quantities at interfaces (maintaining exact conservation, and eliminating all redundant pair-wise interations; see Springel 2010).

(v) We adopt adaptive gravitational softenings as described above. This imposes essentially no cost for gas (since h_i must be computed already for hydrodynamics), but allows us to take much larger timesteps for low-density particles, and (more important for this problem) avoid over-softening particles in dense regions (where a too-large softening radius might encompass thousands of neighbor particles, which imposes a substantial cost in the tree-gravity calculation).

(vi) We have developed and use a hybrid OpenMP-MPI parallelization of the code, which allows us to extend the weak scaling of the code considerably further as we go to large processor

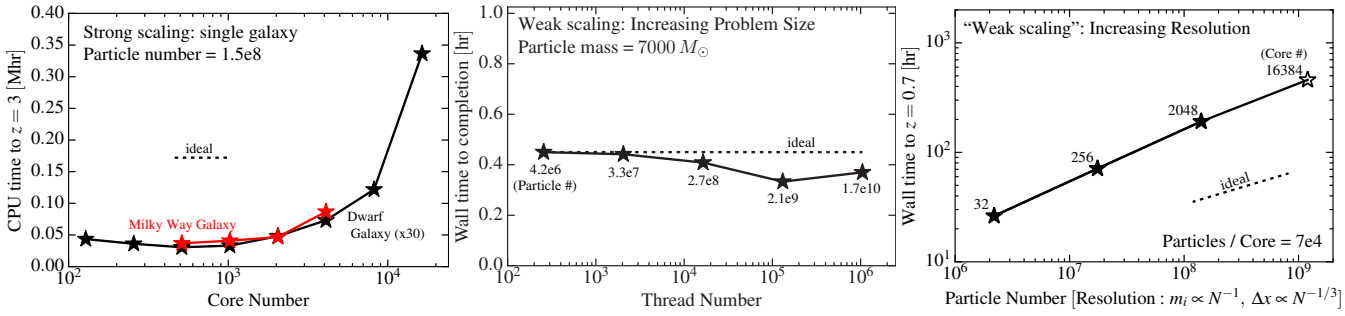

Figure G1. Code scalings of GIZMO in full production-quality FIRE-2 simulations, at our production resolution, with the full, identical physics of gravity, hydrodynamics, cooling, star formation, and feedback to our published simulations. *Left:* Strong scaling for a zoom-in of a MW-mass (**m12i**) or dwarf galaxy (**m10q**) halo, each using 1.5×10^8 particles, run to 25% of the age of the Universe, using the optimal MPI+OpenMP hybrid configuration at each core number. Our optimizations allow us to maintain near-ideal strong scaling to $\sim14,000$ cores per billion particles (2048 for the problem shown). *Center:* Weak scaling, for a full cosmological box, populated with the same high-resolution particles, run for a short fraction of the age of the Universe. Here, we keep resolution fixed at baryonic particle mass $7000M_\odot$, identical to our high-resolution MW simulation at *left*, but we increase the cosmological volume from 2 to 10^4 comoving Mpc^3. The weak scaling of GIZMO's gravity+MHD algorithm is near-ideal (actually slightly *better* at intermediate volume, owing to fixed overheads and increasing statistical homogeneity of the box at larger sizes), to greater than a million threads (here 2^{20} threads, 2^{18} cores, 2^{17} MPI tasks, 2^{14} nodes). *Right:* "Weak scaling" test, increasing the resolution instead of the problem size (specifically, increasing the particle number for the same MW-mass galaxy). Because the resolution increases (hence timestep decreases) with particle number here, the ideal weak scaling for a converged solution is wall-clock time $\propto t_{\rm Hubble}/\Delta t\propto(\Delta x/c_s)^{-1}\propto N^{1/3}$, shown. Our achieved scaling is only slightly worse than this, because new, dense structures such as star clusters are resolved at higher resolution.

number (where we previously found communication costs between neighbors, which are alleviated by the shared-memory structure of OpenMP, were beginning to dominate the run time).

(vii) Our MFM hydrodynamic solver is actually slightly faster than the P-SPH algorithm used for FIRE-1. Hopkins (2015) compare run-times on three problems: a 3D Kelvin-Helmholtz instability, an isolated (non-cosmological) galaxy with star formation and feedback, and the cosmological but strictly ideal-gas (no star formation or cooling) Santa Barbara cluster test. They show that compared to the P-SPH formulation of SPH from Hopkins (2013a) (which incorporated improvements in artificial diffusion terms as well as a larger neighbor number needed in SPH to capture certain instabilities), the speedup on these tests ranged from a factor $\sim1.3-2.5$, mostly owing to the larger neighbor number needed in P-SPH to achieve comparable accuracy.

(viii) We also stress that the inclusion of realistic feedback itself greatly speeds up the code (for galaxy formation simulations) – perhaps more than any purely numerical optimization. Dense regions which would otherwise slow down the computation (as described above) tend to be quickly destroyed by stellar feedback. Without feedback, it would be impossible to run the simulations here with the same resolution and cooling physics below redshift $z\sim2$, because the extremely dense relic star clusters would require constant, extremely short timesteps.

Readers interested in more details should consult the public source code.

Figure G1 demonstrates the scaling of the code GIZMO on a production quality set of FIRE-2 simulations, including all the physics of our production runs (full cosmological integration with self-gravity, baryonic physics including cooling, star formation, and stellar feedback, etc.). This is a "real world" test, as opposed to the scaling on idealized test problems (which can, of course, be much better). All runs were run with an otherwise identical version of the code; at each CPU number the optimal OpenMP-MPI configuration was used. Strong and weak-resolution scaling tests were run on the XSEDE Stampede machine, weak-problem size runs on the DOE ALCF Mira machine.

Our optimizations allow us to extend good strong scaling, at our modest "typical FIRE-1" resolution, to ~1024 cores. Even more strikingly, the optimizations we have made allow us to maintain good weak scaling up to ~4096 cores for a simulation with 3×10^8 particles – our "Latte" resolution – and $\sim16,384$ cores for a zoom-in simulation with $>10^9$ particles (and as many as $\sim10^6$ cores for large-volume simulations with $\sim10^{10}$ particles, which naturally exhibit superior weak scaling). This is especially non-trivial for these sorts of problems, since the gravitational softening is adaptive, so higher particle number implies smaller force softening and hence smaller timesteps in the dense regions.

For comparison, if we compare similar runs to the strong-scaling tests in Fig. G1 with the public GADGET-2 code (using the simpler Springel et al. (2005) sub-grid model for ISM physics and feedback, with wind mass-loading set by-hand to produce a similar mass as our GIZMO runs), we find the scaling saturates at $\sim64-128$ cores.

Figure H1. Believeland.

APPENDIX H: CONGRATULATIONS!

You made it to the end of this paper. Fig. H1 is your reward. Go Cleveland!