

Tests of subgrid models for star formation using simulations of isolated disk galaxies

Folkert S. J. Nobels¹*, Joop Schaye¹, Matthieu Schaller^{1,2}, Sylvia Ploeckinger^{2,3}, Evgenii Chaikin¹, and Alexander J. Richings^{4,5,6}

¹Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

²Lorentz Institute for Theoretical Physics, Leiden University, PO Box 9506, NL-2300 RA Leiden, The Netherlands

³Department of Astrophysics, University of Vienna, Türkenschanzstrasse 17, 1180 Vienna, Austria

⁴Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

⁵E. A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Hull, HU6 7RX, UK

⁶DAIM, University of Hull, Cottingham Road, Hull, HU6 7RX, UK

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We use smoothed-particle hydrodynamics simulations of isolated Milky Way-mass disk galaxies that include cold, interstellar gas to test subgrid prescriptions for star formation (SF). Our fiducial model combines a Schmidt law with a gravitational instability criterion, but we also test density thresholds and temperature ceilings. While SF histories are insensitive to the prescription for SF, the Kennicutt-Schmidt (KS) relations between SF rate and gas surface density can discriminate between models. We show that our fiducial model, with an SF efficiency per free-fall time of 1 per cent, agrees with spatially-resolved and azimuthally-averaged observed KS relations for neutral, atomic and molecular gas. Density thresholds do not perform as well. While temperature ceilings selecting cold, molecular gas can match the data for galaxies with solar metallicity, they are unsuitable for very low-metallicity gas and hence for cosmological simulations. We argue that SF criteria should be applied at the resolution limit rather than at a fixed physical scale, which means that we should aim for numerical convergence of observables rather than of the properties of gas labelled as star-forming. Our fiducial model yields good convergence when the mass resolution is varied by nearly 4 orders of magnitude, with the exception of the spatially-resolved molecular KS relation at low surface densities. For the gravitational instability criterion, we quantify the impact on the KS relations of gravitational softening, the SF efficiency, and the strength of supernova feedback, as well as of observable parameters such as the inclusion of ionized gas, the averaging scale, and the metallicity.

Key words: galaxies: general – galaxies: evolution – galaxies: star formation – ISM: evolution – ISM: structure – methods: numerical

1 INTRODUCTION

Cosmological hydrodynamical simulations of representative volumes typically only resolve scales of $\sim 0.1 - 1$ kpc and masses of $\sim 10^4 - 10^6 M_\odot$ (e.g. Schaye et al. 2015; Pillepich et al. 2018; Nelson et al. 2019; Davé et al. 2019; Dubois et al. 2021; Feldmann et al. 2023). Because much of the physics of star formation (SF) remains unresolved, idealised SF laws are used to determine the SF rates (SFRs) of individual resolution elements. Generally, a so-called Schmidt (1959) law is adopted, which assumes that gas collapses on the density-dependent free-fall time, $t_{\text{ff}} = \sqrt{3\pi/32G\rho}$, and is converted into stars with a specified (fractional) SF efficiency

(SFE) per free-fall time, ε ,

$$\dot{\rho}_* = \varepsilon \frac{\rho}{t_{\text{ff}}}. \quad (1)$$

The SFE ε is a free parameter, typically set between 1 per cent (e.g. Semenov et al. 2017) and 100 per cent (e.g. Hopkins et al. 2018).

Alternative approaches can however also be found in the literature. For example, the Schaye & Dalla Vecchia (2008) subgrid model uses the gas pressure to set the SFR of a resolution element. For simulations that do not resolve the multiphase interstellar medium (ISM) advantages of this approach are that, unlike the density, the pressure does not change dramatically between simulations that do and do not resolve a cold (molecular) gas phase, and that for a self-gravitating disk the pressure is closely related to the gas surface density that appears in the observed kpc-scale Kennicutt (1998) SF surface density law. However, these features are arguably undesirable for high-resolution models that capture the physics resulting in

* E-mail: nobels@strw.leidenuniv.nl

a multiphase ISM and which can be tested by comparing observed and predicted coarse-grained SFEs. Other subgrid prescriptions use physically-motivated SFE models that depend not only on the density, but also on the velocity dispersion (e.g. [Semenov et al. 2016](#); [Kretschmer & Teyssier 2020](#)).

Applying an SF law to all gas may be unrealistic given that stars are observed to form in cold and dense gas. Moreover, we do not wish to impose subgrid models on resolved scales. Therefore, a criterion needs to be applied to determine which gas in the simulation is eligible for SF. Typically, the selected gas satisfies one (or multiple) of the following criteria: (i) the physical gas density exceeds a hydrogen number density threshold $n_{\text{H,crit}}$ that is either constant (e.g. [Springel & Hernquist 2003](#)) or depends on metallicity ([Schaye 2004](#)), (ii) is colder than a critical temperature T_{crit} (e.g. [Wang et al. 2015](#); [Revaz & Jablonka 2018](#)), (iii) is Jeans (1902) unstable (e.g. [Stinson et al. 2006](#); [Hopkins et al. 2018](#)), (iv) is in a converging flow (e.g. [Stinson et al. 2006](#); [Hopkins et al. 2023](#)) and/or (v) is gravitationally bound (e.g. [Hopkins et al. 2014](#)). Additionally, cosmological simulations often include a criterion that only allows gas to be star-forming if it is $\sim 10^2$ times denser than the cosmic mean (e.g. [Schaye et al. 2015](#)). This prevents SF in the gas outside of haloes at very high redshifts where the critical physical density threshold becomes comparable to the cosmic mean.

Observationally, galaxy-averaged SFR surface density relations were shown by [Kennicutt \(1989, 1998\)](#) to follow

$$\Sigma_{\text{SFR}} = A \Sigma_{\text{H+H}_2}^N, \quad (2)$$

where Σ_{SFR} is the SFR per unit area, $\Sigma_{\text{H+H}_2}$ is the neutral hydrogen surface density, A is the normalisation and $N \approx 1.4$ is the slope of the power law. [Kennicutt et al. \(2007\)](#) showed that spatially-resolved kpc-sized regions in individual galaxies follow the same surface density law, which is often referred to as the Kennicutt-Schmidt (KS) relation. Over the last decade, observational surveys have measured azimuthally averaged and spatially resolved scaling relations for large collections of galaxies on spatial scales of ~ 750 pc, thus providing valuable constraints on models of galaxy and SF (e.g. [Sánchez et al. 2012](#); [Bundy et al. 2015](#); [Fogarty et al. 2015](#); [Leroy et al. 2021](#); [Lin et al. 2019](#)). These include atomic KS relations, $\Sigma_{\text{SFR}} = A \Sigma_{\text{H I}}^N$ (e.g. [Bigiel et al. 2008, 2010](#); [Schruba et al. 2011](#)) and molecular KS relations, $\Sigma_{\text{SFR}} = A \Sigma_{\text{H}_2}^N$ (e.g. [Bigiel et al. 2008, 2010](#); [Onodera et al. 2010](#); [Schruba et al. 2011](#); [Bolatto et al. 2017](#); [Lin et al. 2019](#); [Ellison et al. 2020](#); [Pessa et al. 2021](#); [Querejeta et al. 2021](#); [Abdurro'uf et al. 2022](#)). The observations indicate that the azimuthally-averaged SFR in disk galaxies quickly decreases beyond the optical radius (e.g. [Kennicutt 1989](#); [Martin & Kennicutt 2001](#); [Koopmann & Kenney 2004](#); [Bigiel et al. 2010](#)), while the gas density does not decrease as fast because the H I disk typically extends far beyond the optical disk (e.g. [Bosma 1981](#); [Broeils & van Woerden 1994](#); [Reeves et al. 2015](#)). As a result of this sudden drop in SF, a ‘break’ is seen in the azimuthally-averaged and spatially-resolved total gas KS relation around surface densities of $\Sigma_{\text{H+H}_2} = 10 M_{\odot} \text{ pc}^{-2}$ (e.g. [Martin & Kennicutt 2001](#); [Bigiel et al. 2008, 2010](#); [Dessauges-Zavadsky et al. 2014](#)). The presence of a break in the KS relation is one of the strongest indications that simulations require an SF criterion.

CO observations of nearby galaxies show that on scales of $\sim 10^2$ pc molecular gas with a velocity dispersion around 14 km s^{-1} has almost a factor three longer molecular gas depletion time¹ than

gas with velocity dispersions of 12 km s^{-1} ([Leroy et al. 2017](#)). Similarly, spatially-resolved (10^2 pc) observations that target denser gas (traced by HCN and/or HCO^+) show that a higher velocity dispersion correlates with a longer molecular gas depletion time scale (e.g. [Murphy et al. 2015](#); [Viaene et al. 2018](#); [Querejeta et al. 2019](#)). Furthermore, observations of individual molecular clouds show that the observed depletion times are spread over around two orders of magnitude (e.g. [Heiderman et al. 2010](#); [Murray 2011](#); [Evans et al. 2014](#); [Lee et al. 2016](#); [Vutisalchavakul et al. 2016](#); [Ochsendorf et al. 2017](#); [Pokhrel et al. 2021](#)) and that more massive giant molecular clouds (GMCs), which according to [Larson’s \(1981\)](#) empirical laws, have higher velocity dispersions, have longer depletion time scales (e.g. [Lee et al. 2016](#); [Ochsendorf et al. 2017](#)). These observations hence indicate that on scales of 100 pc, gas with a higher velocity dispersion is less likely to form stars. This suggests that an SF criterion based on the velocity dispersion or gravitational instability of gas might be in better agreement with observations than e.g. a constant density threshold.

Numerical simulations have been performed to investigate the diversity of the depletion time in GMCs. They found that the gas depletion time in GMCs varies over more than one order of magnitude (e.g. [Padoan et al. 2012](#); [Grisdale et al. 2019](#); [Grudić et al. 2019](#)), that the average GMC depletion time scale is consistent with an SFE of 1 per cent (e.g. [Grisdale et al. 2019](#); [Grudić et al. 2019](#)) (though some theoretical works predict higher SFEs; e.g. [Raskutti et al. 2016](#)). In addition, they predict that the diversity of the SFE can be caused by several factors such as the gravitational boundedness of the GMCs ([Padoan et al. 2012](#)), the large variations in the cloud properties ([Grisdale et al. 2019](#)), and the mass of the GMC ([Grisdale 2021](#)). Simulations of galaxies have been performed using gas depletion times that depend on the boundedness of clouds ([Semenov et al. 2016](#); [Gensior et al. 2020](#); [Kretschmer & Teyssier 2020](#)), or simply using an SF criterion based on the boundedness of the cloud (e.g. [Hopkins et al. 2013, 2014, 2018, 2023](#); [Semenov et al. 2017, 2018, 2019](#)). Furthermore, simulations of dwarf galaxies have been performed that look at different spatially-averaging scales ([Hu et al. 2016](#)) and splitting the KS relations by hydrogen species ([Whitworth et al. 2022](#)). Not much attention has however been given to comparing them to KS relations split by hydrogen species or different spatially-averaging approaches in Milky Way mass galaxies.

To test subgrid models for SF, simulations of idealised, isolated galaxies offer some advantages over full, cosmological simulations. They offer more control, are less computationally expensive, and do not suffer from the chaotic behavior due to e.g. slight changes in the merger history that complicated comparisons of individual objects in different cosmological simulations using identical initial conditions ([Genel et al. 2019](#); [Keller et al. 2019](#); [Borrow et al. 2023](#)). These features make such simulations well suited for systematic explorations of parameter space and convergence tests.

In this work, we use hydrodynamical simulations of isolated disk galaxies that include a cold, molecular gas phase to systematically investigate several criteria for SF. For our fiducial subgrid model, which consists of a Schmidt law combined with a gravitational instability criterion, we compare with observations of both azimuthally averaged and spatially resolved KS laws for neutral, atomic and molecular gas. We systematically investigate the effect

component under consideration: neutral hydrogen, molecular hydrogen or atomic hydrogen.

¹ We define the depletion time as $t_{\text{dep},i} = \Sigma_i / \Sigma_{\text{SFR}}$, where i is the gas

of the choice of SF criterion, the parameter values, and the numerical resolution.

The remainder of this paper is structured as follows. In Section 2 we describe our subgrid model for ingredients other than the SF criterion and our disk galaxy set-up. This is followed by an overview of the different criteria for SF that we compare: a density threshold, a temperature ceiling, and a gravitational instability criterion. In Section 4 we show our results and in Section 5 we discuss the effect of gravitational softening and compare our results with previous work. In Section 6 we summarise our conclusions.

2 SIMULATIONS

We perform hydrodynamical simulations of isolated disk galaxies using the publically available code SWIFT (Schaller et al. 2016; Schaller et al. 2018, 2023). We use a fast multipole method (Green-gard & Rokhlin 1987) as the gravity solver together with a static external Hernquist (1990) potential. For the hydrodynamics, we use the SPHENIX scheme and parameter values of Borrow et al. (2022). SPHENIX is a density-energy smoothed-particle hydrodynamics (SPH) scheme designed to capture shocks and contact discontinuities with artificial viscosity following Cullen & Dehnen (2010) and artificial conduction following Price (2012). We use a quartic spline for the SPH kernel with the mean weighted number of neighbours given by $\langle N_{\text{ngb}} \rangle \approx 65$. The time steps of particles are limited by the local acceleration ($\Delta t \propto 1/\sqrt{|\mathbf{a}|}$) and limited to 1 per cent of the circular orbital period at the particle's position (similar to what was done in Nobels et al. 2022). Additionally, for gas, the time steps are limited by the Courant-Friedrichs-Lewy (CFL) condition ($C_{\text{CFL}} = 0.2$) and the Durier & Dalla Vecchia (2012) time step limiter is used to prevent vastly different time steps between neighbouring gas particles. We use a fiducial baryonic resolution of $10^5 M_{\odot}$, with a corresponding gravitational Plummer-equivalent softening length of 200 pc. Lastly, we limit the smoothing length to a minimum of $1.55 \times 2 \text{ pc} = 3.1 \text{ pc}$ to prevent artificial collapse (see Ploeckinger et al. 2024) and a maximum of 10 kpc.

2.1 Radiative cooling

Radiative cooling and heating rates are calculated using the non-equilibrium chemistry network Chimes (Richings et al. 2014a,b) for hydrogen and helium species and free electrons. The radiation field consists of a redshift-dependent metagalactic UV background (Faucher-Giguère 2020) and an interstellar radiation field whose intensity increases with the Jeans column density to the power of 1.4, following the KS relation (see Ploeckinger & Schaye 2020, Ploeckinger et al. in prep. for details). Radiation is shielded by a Jeans column density that accounts for both a thermal and a constant turbulent (velocity dispersion of 6 km s^{-1}) pressure component. The cosmic ray rate follows the same scaling as the interstellar radiation field for low column densities but saturates at a cosmic ray rate of $2 \times 10^{-16} \text{ s}^{-1}$ for $N_{\text{H}} \geq 10^{21} \text{ cm}^{-2}$ (Indriolo et al. 2015). In order to account for unresolved clumping, the reactions on the surface of dust grains, such as the formation of H_2 , are boosted above 0.1 cm^{-3} as $B_{\text{dust}} = 10^{1/3} (n_{\text{H}}/1 \text{ cm}^{-3})^{1/3}$, the boost factor saturates at $B_{\text{dust}} = 10$ for densities $n_{\text{H}} \geq 100 \text{ cm}^{-3}$.

The cooling and heating rates of C, N, O, Ne, Mg, Si, S, Ca, and Fe are pre-tabulated using the same radiation field and shielding length assumptions as for the non-equilibrium rates of hydrogen and helium. Their ion fractions are calculated using Chimes assuming

chemical equilibrium and the rates are calculated based on the individual species fractions. A dust-to-metal ratio of $5.6 \times 10^{-3} Z/Z_{\odot}$ is assumed for neutral gas with a cut-off towards higher temperatures ($T > 10^5 \text{ K}$) for which thermal sputtering destroys dust grains on short timescales. The complete set of cooling tables will be made public with a forthcoming publication (Ploeckinger et al. in prep.).

2.2 Star formation rates

When gas is selected to be star-forming (see §3 for a description of the different SF criteria that we employ), it is assigned an SFR following a Schmidt (1959) law (equation 1). We use a constant SFE $\varepsilon = 0.01$, which is motivated by observations of GMCs in the Milky Way (e.g. Vutisalchavakul et al. 2016; Lee et al. 2016; Pokhrel et al. 2021; Hu et al. 2022), the Large Magellanic Cloud (e.g. Ochsendorf et al. 2017), and other nearby galaxies (e.g. Utomo et al. 2018). In § 4.2.2 we investigate the impact that the SFE has on the observed KS relations.

Because the gas consumption time scale $\rho/\dot{\rho}_{\star} = t_{\text{ff}}/\varepsilon$ is much longer than the typical time step size, the Schmidt law needs to be implemented stochastically and the probability of a gas particle converting to a star particle during a time step Δt is:

$$\text{Prob} = \min\left(\frac{\dot{\rho}_{\star} \Delta t}{\rho}, 1\right) = \min\left(\frac{\varepsilon \Delta t}{t_{\text{ff}}}, 1\right). \quad (3)$$

2.3 Stellar feedback

After gas particles have been stochastically converted into stellar particles with mass m_{\star} , they are assumed to represent simple stellar populations that follow a Chabrier (2003) initial mass function (IMF) with a zero age main sequence mass range of $0.1 - 100 M_{\odot}$. We turn off chemical enrichment to ensure that the simulations have a constant metallicity similar to the observations.

2.3.1 Early stellar feedback

We account for three pre-supernova feedback processes: H II regions, stellar winds, and radiation pressure. The implementation of early stellar feedback, whose effects are modest compared to those of supernova feedback, will be described in detail by Ploeckinger et al. (in prep.). We will only give a short explanation of each process below.

The total mass of the H II regions is based on the expected size of the Strömgren (1939) sphere given the local density and the average flux of hydrogen ionizing photons per unit stellar mass during the time step, which is obtained from the Binary Population and Spectral Synthesis (BPASS) stellar evolution and spectral synthesis models (Eldridge et al. 2017; Stanway & Eldridge 2018). Gas in H II regions is assumed to be ionized, non star-forming, and its temperature is limited to $\geq 10^4 \text{ K}$. The radiation pressure is computed using the BPASS photon energy spectrum and the wavelength-dependent optical depth $\tau(\lambda)$ implied by the shielding column used for the cooling tables of Ploeckinger & Schaye (2020) and the tables used here, which scales with the local Jeans length. The momentum from stellar winds is taken from the BPASS tables. The momentum injected during a time step by radiation pressure and stellar winds from a star particle is injected by stochastically kicking neighbouring gas particles with a velocity $v_{\text{kick}} = 50 \text{ km s}^{-1}$ radially away from the stellar particle.

2.3.2 Supernova feedback

We assume that stars with initial mass $> 8 M_{\odot}$ produce core-collapse supernovae (CC SNe), which for our IMF corresponds to $1.18 \times 10^{-2} M_{\odot}^{-1}$ CC SN per unit stellar mass formed. Each CC SN provides a total energy of $E_{\text{CCSN}} = 2.0 \times 10^{51}$ erg. We choose the same value as we used in [Chaikin et al. \(2023\)](#). This value is slightly higher than the canonical 10^{51} erg, but may be viewed as accounting for the ~ 10 times more energetic hypernovae, stars more massive than $100 M_{\odot}$, and/or simply as compensation for some degree of numerical overcooling. We use metallicity-dependent stellar lifetime tables to calculate the number of CC SN that explode every time step (based on [Portinari et al. 1998](#)). Following [Chaikin et al. \(2023\)](#), star particles inject 90 per cent of their CC SN energy in thermal form and the remaining 10 per cent as kinetic energy. The thermal part is injected using the stochastic method of [Dalla Vecchia & Schaye \(2012\)](#) using the statistically isotropic selection of gas particles of [Chaikin et al. \(2022\)](#). The temperature of the gas receiving thermal CC SN feedback is increased by $\Delta T = 10^{7.5}$ K to suppress numerical overcooling. The remaining 10 per cent of the energy is injected kinetically using the energy, momentum and angular momentum conserving method detailed in [Chaikin et al. \(2023\)](#). Pairs of particles are selected isotropically and kicked in opposite directions with a desired velocity of $\Delta v_{\text{CCSN}} = 50 \text{ km s}^{-1}$ (the actual kick velocity depends on the relative velocity of the star and gas particles). The high-energy thermal events drive galactic winds and can be thought of as representing superbubbles due to clustered CC SN events, while the low-energy kinetic events drive turbulence and can be viewed as representing the momentum-driven phase of isolated CC SN. [Chaikin et al. \(2023\)](#) found that their prescription and fiducial parameter values, which we adopt here, led to good agreement with the observed relation between the gas velocity dispersion and the star formation rate surface density. For more details and discussion we refer to [Chaikin et al. \(2023\)](#).

For type-Ia supernova (SNIa) feedback, the exact delay of SNIa cannot be predicted because it depends on poorly constrained parameters (i.e. the binary fraction and separation). Therefore, we use a statistical approach that samples the SNIa rate from a delay time distribution (DTD) as

$$\text{DTD}(t) = \frac{\nu}{\tau} e^{-(t-t_{\text{delay}})/\tau} \Theta(t-t_{\text{delay}}), \quad (4)$$

where $\nu = 2.0 \times 10^{-3} M_{\odot}^{-1}$ is the number of SNIa per unit formed stellar mass ([Maoz & Mannucci 2012](#)), $\tau = 2 \text{ Gyr}$ (based on Fig 2 of [Maoz et al. 2010](#)), $\Theta(x)$ is the Heaviside step function and $t_{\text{delay}} = 40 \text{ Myr}$ corresponds to the maximum lifetime of a star that explodes as a CC SN ([Portinari et al. 1998](#)). SNIa energy is injected 100 per cent thermally and isotropically as described above for CC SN, but using 10^{51} erg per SNIa.

2.4 Initial conditions

We simulate an isolated galaxy consisting of an exponential disk of stars and gas embedded in a dark matter halo. To reduce unnecessary computational expense, the dark matter halo is modelled using an external gravitational potential instead of with dark matter particles.

The initial conditions are generated with the `MakeNewDisk` code ([Springel et al. 2005](#); hereafter [S05](#)). The dark matter halo follows a [Hernquist \(1990\)](#) profile, but compared to [S05](#) we use a slightly different absolute normalisation and a scale radius that produces a better match to a dark matter halo that follows a [Navarro et al. \(1997\)](#) (hereafter [NFW](#)) profile. We define the [Hernquist \(1990\)](#)

profile as:

$$\rho_{\text{Hern}}(r) = \frac{M_{\text{Hern}}}{2\pi} \frac{r_{\star}}{r(r+r_{\star})^3}, \quad (5)$$

where M_{Hern} is the total mass obtained when integrating the profile to $r = \infty$ and r_{\star} is the scale radius (and also the half-mass radius). To set the total halo mass, we use the parameter M_{200} , where M_{200} is the mass within R_{200} , the radius within which the average density of the halo is 200 times the critical density of the universe. Using $M_{200} = \int_0^{R_{200}} \rho_{\text{Hern}} 4\pi r^2 dr$ gives

$$M_{200} = \frac{M_{\text{Hern}} R_{200}^2}{(R_{200} + r_{\star})^2}. \quad (6)$$

This means the [Hernquist \(1990\)](#) profile can be written as

$$\rho_{\text{Hern}}(r) = \frac{M_{200}(R_{200} + r_{\star})^2}{2\pi R_{200}^2} \frac{r_{\star}}{r(r+r_{\star})^3}. \quad (7)$$

When we constrain the central profile to follow the [NFW](#) profile

$$\rho_{\text{NFW}}(r) = \frac{M_{200}}{\frac{4\pi}{3} R_{200}^3} \frac{1}{3 \left(\ln(1+c) - \frac{c}{1+c} \right) \left(\frac{r}{R_{200}} \left(\frac{r_{\star}}{R_{200}} + \frac{r}{R_{200}} \right) \right)^2}, \quad (8)$$

we obtain:

$$\left(\frac{r_{\star}}{R_{200}} \right)^2 = \frac{2}{c^2} \left(\ln(1+c) - \frac{c}{1+c} \right) \left(\frac{r_{\star}}{R_{200}} + 1 \right)^2, \quad (9)$$

where c is the [NFW](#) concentration parameter. [S05](#) made the simplifying assumption that $r_{\star}/R_{200} \ll 1$ and omitted the factor $(r_{\star}/R_{200} + 1)^2$ from equation (9). However, in order to obtain a better match to the [NFW](#) profile this assumption is not required. The only positive root of equation (9) for the scale length is

$$r_{\star} = \frac{b + \sqrt{b}}{1-b} R_{200}, \quad (10)$$

where b is defined as

$$b \equiv \frac{2}{c^2} \left(\ln(1+c) - \frac{c}{1+c} \right), \quad (11)$$

and $R_{200} = (GM_{200}/100H_0^2)^{1/3}$, where G is the gravitational constant and $H_0 = 70.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is the Hubble constant. In the limit $\sqrt{b} \ll 1$ (which corresponds to large [NFW](#) concentrations), we have $b \ll \sqrt{b}$ so that $r_{\star} \ll R_{200}$, and the result of [S05](#) is obtained. The calculation of the density distributions and the velocities of the stellar and gas particles remain the same as in [S05](#), but with a differently normalised [Hernquist \(1990\)](#) profile. We set the angular momentum J of the dark matter halo using the dimensionless spin parameter $\lambda \equiv J|E|^{1/2}/GM_{200}^{5/2}$, where E is the total energy of the dark matter halo.

We choose a virial mass $M_{200} = 1.37 \times 10^{12} M_{\odot}$ and a concentration parameter $c = 9$ (based on the mass-concentration relation from [Correa et al. 2015](#), using the [Planck Collaboration et al. \(2016\)](#) cosmology). The dimensionless spin parameter of the halo is set to $\lambda = 0.033$ ([Oppenheimer 2018](#)), which yields a radial disk scale length of $r_d = 4.3 \text{ kpc}$ ([Mo et al. 1998](#); [S05](#)). The disk mass is $M_{\text{disk}} = 5.48 \times 10^{10} M_{\odot}$, which corresponds to 4 per cent of the virial mass, of which the stellar disk contains 70 per cent, $M_{\text{disk},\star} = 3.836 \times 10^{10} M_{\odot}$, with a fixed stellar disk scale-height of 10 per cent of the radial disk scale length, i.e. 0.43 kpc . The remaining 30 per cent of the disk is gas, corresponding to $M_{\text{disk,gas}} = 1.644 \times 10^{10} M_{\odot}$. Using these parameter values the stellar mass is comparable to those of the more massive disk galaxies in observations of KS relations (e.g. [Bigiel et al. 2008, 2010](#)) and

we sample gas surface densities up to $\sim 100 M_{\odot} \text{ pc}^{-2}$ as observed for massive disk galaxies (e.g. [Walter et al. 2008](#)). In our fiducial model we assume solar metallicity ($Z = Z_{\odot} = 0.0134$; [Asplund et al. 2009](#)), but we investigate the effect of different metallicities in §4.1.

The gas initially has a temperature of 10^4 K. We assume vertical hydrostatic equilibrium, where the disk scale height of the gas is set by the pressure and the gravity of the gas and stars and therefore depends on the radius. Note that in the centre the initial gas scale height is much smaller than the stellar disk scale height; they only become comparable in the outer region.

To prevent an artificial collapse of the gas disk and a spurious initial burst of SF, we give a fraction of the stellar particles an age distribution corresponding to a constant SFR of $10 M_{\odot} \text{ yr}^{-1}$ over the last 100 Myr and allow them to inject CC SN and SNIa feedback.

The initial conditions are made publicly available as an example in the SWIFT code repository and can be found in the `IsolatedGalaxy` examples at www.swiftsim.com.

3 STAR FORMATION CRITERIA

In this work we investigate the impact of different subgrid prescriptions for SF, in particular different criteria for selecting gas particles that are eligible for conversion into stellar particles. The calculation of the SFRs of star-forming gas particles was described in §2.2. This section describes how the star-forming gas is selected. We will test three different criteria: a density threshold (§3.1), a temperature ceiling (§3.2) and a gravitational instability criterion (§3.3).

3.1 Density threshold

The simplest and most widely used SF criterion is a density threshold,

$$n_{\text{H}} > n_{\text{H,crit}} \quad (12)$$

where $n_{\text{H,crit}}$ is the total hydrogen number density² above which the gas is star-forming.

It is instructive to see what gas phases are selected by criteria with different density thresholds. The dotted curves in Fig. 1 show the equilibrium temperature³ as a function of density for different metallicities. We see that for $Z = Z_{\odot}$ the transition to the cold phase ($T \ll 10^3$ K) occurs at $n_{\text{H}} \sim 10^{-0.5} \text{ cm}^{-3}$ while for very low metallicities (e.g. $Z = 10^{-4} Z_{\odot}$) the equilibrium temperature remains above 10^3 K for densities $n_{\text{H}} \lesssim 10^3 \text{ cm}^{-3}$. This means that for different metallicities the selected gas will correspond to different phases, i.e. at low metallicities the star-forming gas is warm, while at higher metallicities it is cold.

Since observations indicate that SF is associated with cold gas, a constant density threshold has the disadvantage that, depending on the metallicity, the selected gas can be in the warm phase unless the threshold is chosen to be very high. A fixed density threshold is common in the literature (e.g. [Springel & Hernquist 2003](#), and models based on this), while other models use a metallicity-dependent density threshold designed to track the transition from the warm to the cold interstellar gas-phase (e.g. [Schaye 2004](#); [Agertz et al. 2013](#) and models based on these such as EAGLE; [Schaye et al. 2015](#)).

² We use a fixed primordial hydrogen mass fraction $X = 0.756$ to convert the gas mass density into the total hydrogen number density.

³ The equilibrium temperature is the temperature at which the radiative heating rate is equal to the radiative cooling rate.

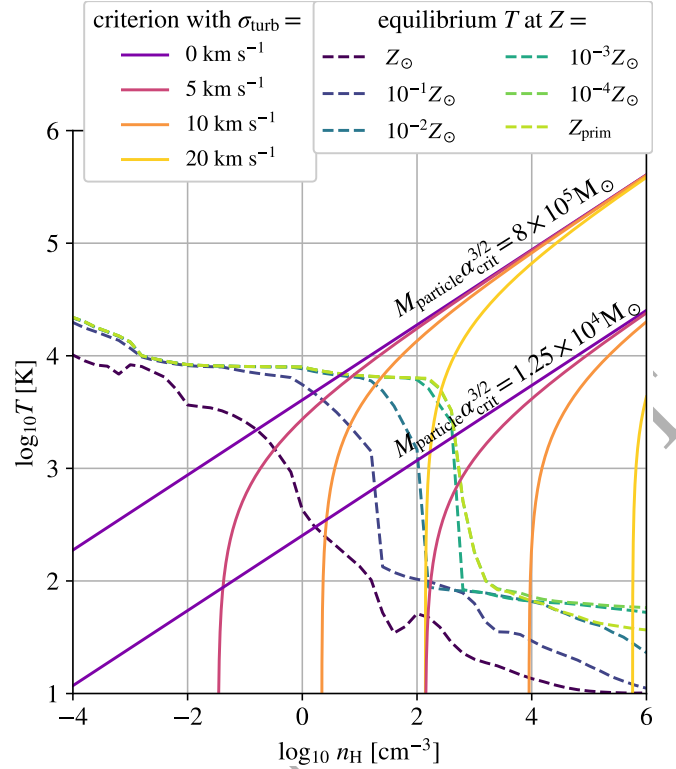


Figure 1. The gravitational instability criterion in temperature-density space for different velocity dispersions (solid lines, different colours; gas at higher densities or lower temperatures is unstable) and different resolutions and/or normalisations (two sets of lines labelled with different values of $M_{\text{particle}} \alpha_{\text{crit}}^{3/2}$). For comparison we also plot the thermal equilibrium temperature curves for different metallicities (dotted lines, different colours) for the Ploekinger et al. (in prep.) cooling tables. The gray lines indicate different density thresholds and different temperature ceilings. Gas with a higher velocity dispersion requires higher densities to be star-forming. The gas selected by the gravitational instability criterion shifts to higher (lower) densities for higher (lower) numerical resolutions, i.e. lower (higher) values of M_{particle} , if the instability criterion (i.e., the value of α_{crit}) is kept constant. Therefore, the gas phase that is selected to be star-forming depends on the numerical resolution, unless we scale $\alpha_{\text{crit}} \propto M_{\text{particle}}^{-2/3}$.

3.2 Temperature ceiling

To restrict the formation of stars to colder gas that resembles the environment where stars are observed to form (e.g. molecular clouds), a temperature criterion is a natural choice,

$$T < T_{\text{crit}} \quad (13)$$

The parameter T_{crit} defines a temperature ceiling, which is usually set to be low enough to only select gas that is expected to have conditions suitable for SF. Looking again at the equilibrium temperature curves in Fig. 1, we see that for very low metallicities a reasonable value like $T_{\text{crit}} = 10^2$ K will only select gas with very high densities, $n_{\text{H}} \gtrsim 10^4 \text{ cm}^{-3}$. If such high densities are not resolved, then the SFR will be underestimated. Cosmological simulations with limited resolution that start with primordial abundances may never form any stars at all. Combining density and temperature criteria, i.e. $n_{\text{H}} > n_{\text{H,crit}}$ or $T < T_{\text{crit}}$, can help alleviate this problem. We will, however, consider a different solution.

3.3 Gravitational instability criterion

Another option for the SF criterion is to select gas that is gravitationally unstable, i.e. that satisfies the [Jeans \(1902\)](#) instability criterion $t_{\text{ff}} < t_{\text{cr}}$, where t_{cr} is the sound and turbulence crossing time which for a cloud of size h is⁴

$$t_{\text{cr}} = \frac{h}{\sqrt{\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2}}, \quad (14)$$

where the thermal dispersion σ_{th} is given by:

$$\sigma_{\text{th}} = \sqrt{3 \frac{P}{\rho}} = \sqrt{\frac{3k_{\text{B}}T}{\mu m_{\text{H}}}} = 13.8 \text{ km s}^{-1} \left(\frac{T}{10^4 \text{ K}} \right)^{1/2}, \quad (15)$$

where we set $\mu = 1.3$ which is appropriate for neutral atomic gas, and σ_{turb} is the turbulent velocity dispersion. In our SPH implementation, the turbulent velocity dispersion is calculated for each particle i and its neighbours j as:

$$\sigma_{\text{turb},i}^2 = \frac{1}{\rho_i} \sum_{j=1}^N M_{\text{particle},j} W(r_{ij}, s_i) |\mathbf{v}_i - \mathbf{v}_j|^2, \quad (16)$$

where M_{particle} is the particle mass, W is the kernel weight function, s is the smoothing length, \mathbf{v}_i and \mathbf{v}_j are the particle velocities, and r_{ij} is the particle separation. We find that this gives identical results to the velocity dispersion calculations of [Hopkins et al. \(2013, 2018\)](#) who compute either a full tensor or both $\nabla \times \mathbf{v}$ and $\nabla \cdot \mathbf{v}$. An advantage of using equation (16) is that only a single SPH variable is required in contrast with 4 or 9 additional SPH variables⁵.

A cloud is gravitationally unstable when its free-fall time is smaller than the sound and turbulence crossing time. We calculate the free-fall time from first principles for consistency with §5.1 where we will investigate the impact of including the gravitational softening length ϵ in the derivation. The free-fall time is given by

$$t_{\text{ff}} = \int_h^0 \frac{dr}{dr}. \quad (17)$$

For an initially static, spherical cloud of mass m and radius h , energy conservation implies that the velocity of the outer shell is given by

$$v(r) = - \left(\frac{2Gm}{r} - \frac{2Gm}{h} \right)^{1/2}. \quad (18)$$

This means that the free-fall time is given by

$$t_{\text{ff}} = \frac{1}{\sqrt{2Gm}} \int_0^h \left(\frac{1}{r} - \frac{1}{h} \right)^{-1/2} dr, \quad (19)$$

$$= \sqrt{\frac{3}{8\pi G\rho}} \int_0^1 \sqrt{\frac{x}{1-x}} dx, \quad (20)$$

$$= \sqrt{\frac{3\pi}{32G\rho}} = 45 \text{ Myr} \left(\frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right)^{-1/2}. \quad (21)$$

⁴ Note that it is possible to include additional terms, for example when MHD is used (e.g. $t_{\text{cr}} = h / \sqrt{\sigma_{\text{th}}^2 + \sigma_{\text{Alfvén}}^2 + \sigma_{\text{turb}}^2}$), where $\sigma_{\text{Alfvén}}$ is the velocity dispersion of ions.

⁵ Six additional variables when symmetry arguments are used.

Equating t_{cr} (equation 14) and t_{ff} (equation 21) gives the generalized [Jeans \(1902\)](#) length:

$$\lambda_{\text{J}} = \sqrt{\frac{3\pi(\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2)}{32G\rho}}, \quad (22)$$

$$= 0.46 \text{ kpc} \left(\frac{\sqrt{\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2}}{10 \text{ km s}^{-1}} \right) \left(\frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right)^{-1/2}, \quad (23)$$

and a corresponding [Jeans \(1902\)](#) mass:

$$M_{\text{J}} = \frac{4\pi}{3} \rho \lambda_{\text{J}}^3 = \frac{4\pi}{3} \left(\frac{3\pi}{32} \right)^{3/2} \left(\frac{\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2}{G} \right)^{3/2} \rho^{-1/2}, \quad (24)$$

$$= 1.3 \times 10^7 M_{\odot} \left(\frac{\sqrt{\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2}}{10 \text{ km s}^{-1}} \right)^3 \left(\frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right)^{-1/2}. \quad (25)$$

We can construct an SF criterion by demanding gravitational instability at the mass resolution limit, i.e., by requiring the [Jeans \(1902\)](#) mass to be smaller than the expected mass within the SPH kernel

$$M_{\text{J}} < M_{\text{res}} = \langle N_{\text{ngb}} \rangle M_{\text{particle}}. \quad (26)$$

Combining equations (24) and (26), we get

$$\alpha \equiv \frac{\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2}{G \langle N_{\text{ngb}} \rangle^{2/3} M_{\text{particle}}^{2/3} \rho^{1/3}} < \alpha_{\text{crit}}, \quad (27)$$

where α_{crit} is a constant of order unity and the gas is considered unstable if $\alpha < \alpha_{\text{crit}}$. Using equation (24) gives $\alpha_{\text{crit}} = 32/3\pi(3/4\pi)^{2/3} \approx 1.3$, but the appropriate value depends on the assumed geometry. Using a smaller values of α_{crit} implies that star forming gas must be more strongly gravitationally bound. Our fiducial value is $\alpha_{\text{crit}} = 1$.

Similar criteria have been used before in hydrodynamical simulations, for example by [Hopkins et al. \(2014, 2018\)](#) and [Semenov et al. \(2016\)](#), and are often called virial criteria. However, our calculation of the gravitational instability criterion is not identical to the ones in those papers. The main difference is that they used the sound speed instead of the thermal dispersion and that they adopted another normalisation. The difference in normalisation is because a gravitational instability criterion requires gas to satisfy $E_{\text{grav}} + E_{\text{kin}} < 0$ while a virial criterion requires gas to satisfy $E_{\text{grav}} + 2E_{\text{kin}} < 0$.

Based on equation (27), it is clear that for a fixed α_{crit} the instability criterion selects different gas for different resolutions M_{particle} . We will therefore investigate two different cases when varying the resolution: a fixed α_{crit} and a scaled $\alpha_{\text{crit}} \propto M_{\text{particle}}^{-2/3}$. In the former case higher gas densities (or lower velocity dispersions) are required when the numerical resolution increases, whereas in the latter case gas with the same physical properties is selected regardless of the numerical resolution.

It is instructive to consider two limiting cases. The first is when the turbulent velocity dispersion is negligible compared to the thermal dispersion (i.e. $\sigma_{\text{th}} \gg \sigma_{\text{turb}}$). In this case, the instability criterion reduces to the thermal Jeans criterion,

$$T < \frac{\mu m_{\text{H}}^{4/3} G}{3k_{\text{B}} X^{1/3}} \alpha_{\text{crit}} \langle N_{\text{ngb}} \rangle^{2/3} M_{\text{particle}}^{2/3} n_{\text{H}}^{1/3}, \quad (28)$$

$$< 2.5 \times 10^3 \text{ K} \left(\frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right) \left(\frac{\langle N_{\text{ngb}} \rangle}{65} \right)^{2/3} \left(\frac{M_{\text{particle}}}{10^5 M_{\odot}} \right)^{2/3} \alpha_{\text{crit}}, \quad (29)$$

where we assumed that the mean molecular mass is $\mu = 1.3$. This version of the instability criterion gives the maximal temperature that star-forming gas can have as a function of the gas density and the numerical resolution. Equation (29) shows that the temperature of star-forming gas depends explicitly on the resolution of the simulation. In Fig. 1, we compare the instability criterion with the equilibrium temperature curves for fixed metallicities. Because of the degeneracy between α_{crit} and M_{particle} , the instability criterion is only shown for fixed $M_{\text{particle}}\alpha_{\text{crit}}^{2/3}$. The figure shows that for a lower (higher) resolution or larger (smaller) α_{crit} the temperature of the selected gas can be higher (lower) for a fixed gas density.

The second limiting case is when the thermal velocity dispersion is negligible compared to the turbulent dispersion. In this case the instability criterion reduces to a density criterion for a fixed velocity dispersion,

$$n_{\text{H}} > \frac{X\sigma_{\text{turb}}^6}{m_{\text{H}}\alpha_{\text{crit}}^3 G^3 \langle N_{\text{ngb}} \rangle^2 M_{\text{particle}}^2}, \quad (30)$$

$$> 0.14 \text{ cm}^{-3} \left(\frac{\sigma_{\text{turb}}}{5 \text{ km s}^{-1}} \right)^6 \left(\frac{\langle N_{\text{ngb}} \rangle}{65} \right)^{-2} \left(\frac{M_{\text{particle}}}{10^5 M_{\odot}} \right)^{-2} \alpha_{\text{crit}}^{-3}. \quad (31)$$

This effective density criterion is very sensitive to the velocity dispersion, which can be seen in Fig. 1 by comparing differently coloured solid curves. This means that gas at a fixed density quickly becomes ineligible for SF as the turbulent velocity dispersion increases. The dependence on the resolution is very strong for simulations with a fixed α_{crit} , but for simulations in which α_{crit} is scaled with the resolution, gas with the same densities is selected.

4 RESULTS

In § 4.1 we will discuss the results of the fiducial model followed, in § 4.2, by variations in the subgrid prescriptions for SF and stellar feedback in order to understand the impact that different aspects of the galaxy formation model have.

4.1 The fiducial model

In this section, we investigate the SF properties of three different simulations that are run using the fiducial model, which uses the gravitational instability criterion for SF (equation 27) with $\alpha_{\text{crit}} = 1$, the Schmidt (1959) law for SF (equation 1) with $\varepsilon = 0.01$, and a particle mass of $10^5 M_{\odot}$. We will compare simulations with metallicities of $0.5Z_{\odot}$, Z_{\odot} , and $2Z_{\odot}$. Before doing so, we first provide a visual impression of the main properties of the galaxy with solar metallicity.

Fig. 2 shows face-on images of different properties of the galaxy in the fiducial simulation with metallicity $Z = Z_{\odot}$ at time $t = 1 \text{ Gyr}$. The top two rows show the surface densities for all gas (left column), atomic hydrogen (middle column), and molecular hydrogen (right column), where the species fractions are calculated using Chimes. The galaxy spiral structures are most prominent in the molecular hydrogen maps, but the atomic hydrogen also shows clear spiral structure. The SFR surface density (third row, right column) mostly traces the molecular hydrogen surface density.

Across most of the galaxy, the thermal and turbulent velocity

dispersions are similar⁶ (compare the left and middle columns second row of the face-on images). The dispersion tends to be small when the surface density is high, and vice versa. However, in the regions with the highest surface densities the mass-weighted virial parameter (bottom left) is still high ($\alpha \gg 1$). Indeed, most of the gas in the disk has a high value of α and is not forming stars (bottom right), while most of the star-forming gas has $\alpha \approx 1$ (bottom centre).

The instability criterion can be converted into a function of the gas surface density Σ and the scale height h . The mass of the cloud is given by $M = \frac{4\pi}{3}\rho h^3$ and the density is given by $\rho = \Sigma/h$. This results in

$$\alpha = \frac{\sigma^2}{GM^{2/3}\rho^{1/3}} = \left(\frac{4\pi}{3} \right)^{-2/3} \frac{\sigma^2}{G\Sigma h}, \quad (32)$$

$$= 7.1 \left(\frac{\Sigma/\sigma^2}{0.4 M_{\odot} \text{ pc}^{-2} (\text{km s}^{-1})^{-2}} \right)^{-1} \left(\frac{h}{40 \text{ pc}} \right)^{-1}. \quad (33)$$

This means that the ratio Σ/σ^2 is a measure of the boundedness of the gas. Leroy et al. (2017) measured Σ/σ^2 for molecular gas on spatial scales from 300 pc to 1 kpc in M51 and their measurements indicate that the stability parameter ranges between $\alpha \approx 5$ and ≈ 14 , implying that most of the molecular gas in the ISM is gravitationally stable. Not surprisingly, these measurements are lower than our mass weighted values that include all gas. The molecular gas mass weighted values are in good agreement with these observations and range between $\alpha \approx 2$ and 20 (not shown).

4.1.1 Star formation histories

Fig. 3 shows the star formation history (SFH) for the simulations with different metallicities. The thick lines show the SFH averaged over ⁷100 Myr, we average the instantaneous SFR that is logged every single time step over 100 Myr. The different simulations are within 0.4 dex of each other and lower metallicity results in a lower SFR, in agreement with the findings of Richings & Schaye (2016) for a different SF criterion. Besides the 100 Myr averaged SFH we also show the instantaneous SFR that is updated every time a gas particle is active and therefore has the same time resolution as the simulation. Here the SFH varies by about 0.2 dex over times scales $\sim 10 \text{ Myr}$. The small artificial initial SF peak at $t < 50 \text{ Myr}$ does not impact the results at $t > 100 \text{ Myr}$ because we include feedback from existing stars. After a few 100 Myr, the SFR declines because the gas in the disk is being ejected and consumed.

4.1.2 The spatially-resolved Kennicutt-Schmidt relation

The SFH in general does not provide a stringent test of models of SF (e.g. Schaye et al. 2010; Hopkins et al. 2013) because it is determined mostly by the gas supply and feedback processes on galaxy scales. Galaxy-averaged KS relations have the same problem because they are effectively SFRs that are re-normalized to the galaxy size. More useful tests are the spatially-resolved Kennicutt-Schmidt (KS) relations (e.g. Kennicutt et al. 2007), between gas surface density and SF,

$$\Sigma_{\text{SFR}} = A \Sigma_i^N, \quad (34)$$

⁶ The colour scale of the gas temperatures is the same as the colour of the thermal velocity dispersion computed for that temperature.

⁷ We average the instantaneous SFR that is logged every single time step over 100 Myr, i.e. $\text{SFR}_{\text{avg}} = \frac{\int_{t'=t-50 \text{ Myr}}^{t'=t+50 \text{ Myr}} \text{SFR}(t') \Delta t(t')}{100 \text{ Myr}}$.

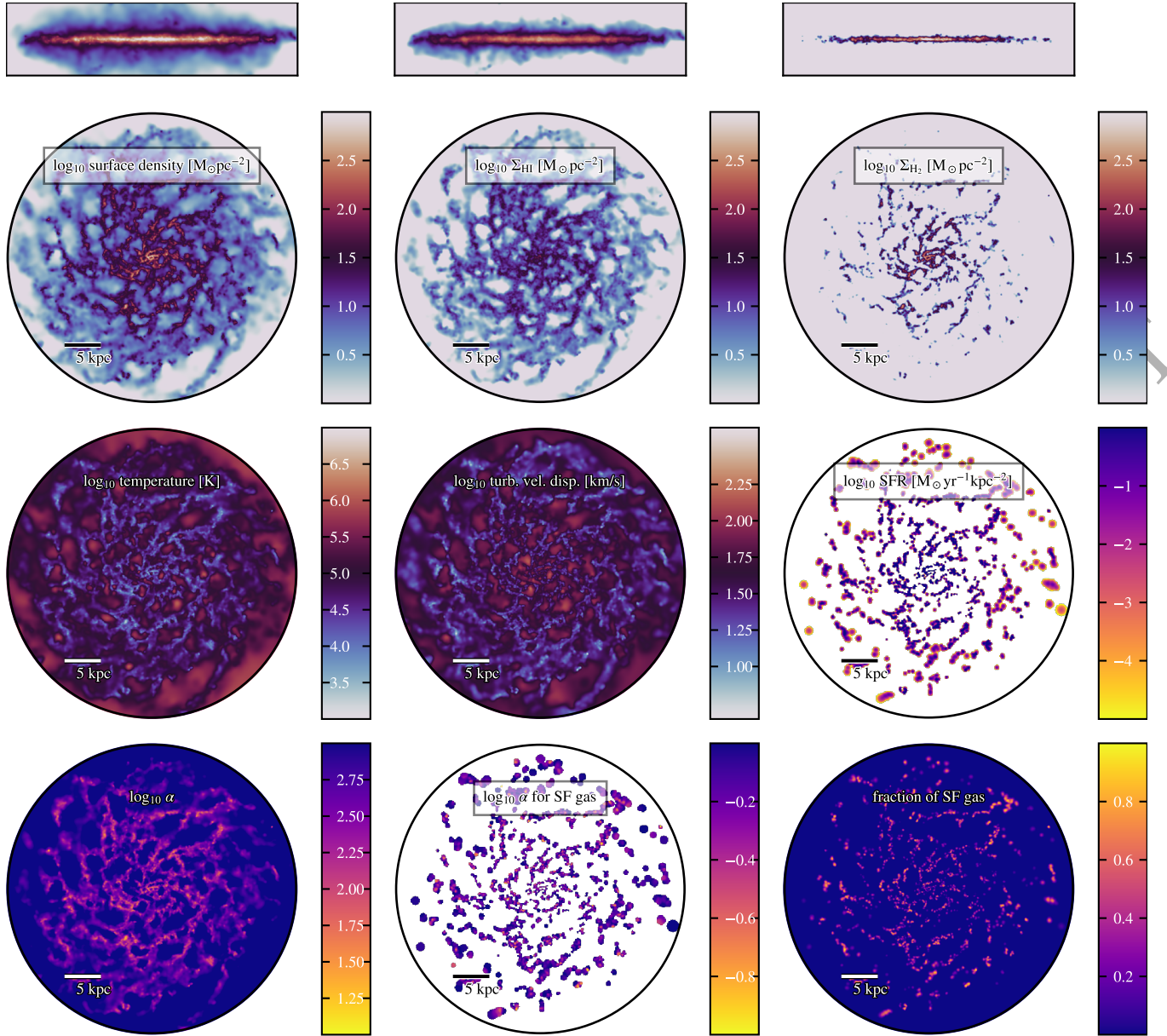


Figure 2. The fiducial simulation with $Z = Z_{\odot}$ at $t = 1$ Gyr. *Top two rows:* Edge-on and face-on gas surface density for the total gas (left column), atomic hydrogen (middle column) and molecular hydrogen (right column). *Third row:* Face-on view of the mass-weighted temperature (left), mass-weighted 3D turbulent velocity dispersion (middle), and SFR surface density (right). *Bottom row:* Face-on view of the mass-weighted virial parameter α for all gas (left) and only star-forming gas (middle), and the mass fraction of gas that is star-forming (right). The thickness of the projection is 50 kpc. The kernel sizes vary across the galaxy from < 0.1 kpc in the centre to around 1 kpc in the outskirts

where Σ_{SFR} is the SFR surface density in a spatially-resolved region, A is the normalisation constant, Σ_i is the surface density of gas in the same region where i is the component (H I , H_2 or $\text{H I} + \text{H}_2$) and N is the slope of the power-law relation.

Focusing again on Fig. 2, we show the total, atomic and molecular gas surface densities for the fid simulation in the top panel. We see that most of the disk area consists of atomic hydrogen. Molecular hydrogen is largely confined to the highest-density regions. Comparing the top right and centre right panels of Fig. 2, we see that the molecular hydrogen surface density is spatially closely correlated with the SFR surface density.

The top three panels of Fig. 4 show the spatially-resolved KS relations for, respectively, neutral, molecular and atomic gas. The dif-

ferent coloured curves show the median relations in 0.25-dex-wide surface density bins for our three different metallicities. The shaded regions indicate the 16th and 84th percentiles for the case of solar metallicity. The relations are averaged over square regions of size 750 pc, matching the typical resolution of observations. In § 4.1.4 we will also consider the azimuthally averaged KS relation (i.e. the KS relation averaged over radial rings of width $\Delta r \approx 750$ pc).

Data points with error bars indicate the median and the 16th-84th percentile scatter for observations of galaxies that have metallicity variations similar to our simulations. None of the relations includes a correction for helium and heavy elements in the surface densities. We compare to observations of global, galaxy disk-averaged KS relations in spiral and star-forming galaxies (Kennicutt

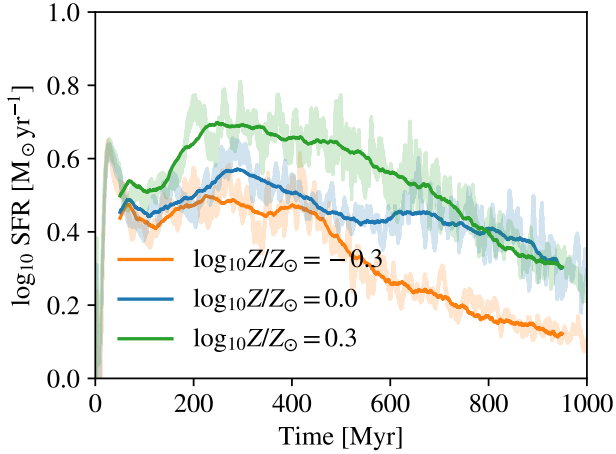


Figure 3. Comparison of the star formation histories for the simulations with different metallicities (different colours). The thick lines show the 100 Myr moving average of the SFR and the transparent lines show the SFR at the time resolution of the simulation. The differences are small but the SFR is higher (lower) for simulations with higher (lower) metallicities. The SFR varies by ≈ 0.2 dex over short time scales.

1998; hereafter K98), of spatially resolved atomic and molecular hydrogen surface densities (Bigiel et al. 2008; hereafter B08; Bigiel et al. 2010; hereafter B10), and of spatially resolved molecular hydrogen surface densities (Ellison et al. 2020; hereafter E20; Pessa et al. 2021; hereafter P21; Querejeta et al. 2021; hereafter Q21). We corrected the surface densities of B08, B10, E20 and Q21 by removing the factor for helium and heavy elements.

We note that there is considerable uncertainty in the observations due to the use of CO as a tracer for molecular hydrogen. First, the α_{CO} factor is assumed to be constant but likely depends on metallicity (e.g. $\propto Z^{-1}$; Narayanan et al. 2012). If gas with higher Σ_i has higher metallicity, then this would imply that the KS relation is steeper (and hence that the t_{gas} relation has a more negative slope). Second, gas with densities below the critical density of CO ($n_{\text{crit,CO}} = 10^3 \text{ cm}^{-3}$; Schöier et al. 2005) is only detected when the clouds have large optical depths τ such that the effective critical density is $n_{\text{crit,eff,CO}} = n_{\text{crit,CO}}/\tau$. This means that CO observations can underestimate the CO mass. If the volumetric density increases with the surface density, then correcting for this effect would again steepen the KS relation. Third, the α_{CO} factor may vary due to differences in the optical depth of the gas caused by higher turbulent velocities (e.g. Teng et al. 2022). Fourth, the α_{CO} factor is smaller inside the centre of galaxies and therefore the slope of the KS relation will steepen (den Brok et al. 2023). Fifth, the commonly adopted value for the ratio of CO (2-1)/(1-0) is 0.65 but observed ratios range from 0.3 to 1.0 (den Brok et al. 2021). There are indications of a trend between this ratio and Σ_{SFR} (den Brok et al. 2021; Yajima et al. 2021). Similarly, the CO (3-2) transition ratios are not constant (Leroy et al. 2022). This might be because of different densities and temperatures in the CO gas, which implies uncertainty in the α_{CO} factor.

In Fig. 4 the agreement between the data and the simulations is good, both for the median and the scatter, and both in terms of the normalization and the slope. The neutral and atomic KS relations agree better with B10 (B08) at low (high) surface densities as would be expected because B10 (B08) observed the outer (inner) parts of the same galaxies. Additionally, below $\Sigma_{\text{H+H}_2} \approx 10 \text{ M}_{\odot} \text{ pc}^{-2}$, B08 detected much fewer regions than B10. The neutral and atomic

KS relations decline sharply below the canonical SF threshold of $\Sigma_i = 10 \text{ M}_{\odot} \text{ pc}^{-2}$ (e.g. Schaye 2004, and references therein). At high $\Sigma_{\text{H+H}_2}$ the slope approaches the canonical value $N \approx 1.4$ (Kennicutt 1998). The slope of the atomic KS relation is very steep, $N \approx 3$, consistent with the observations. For the remainder of this paper, we will use the gas consumption time t_{gas} (bottom of Fig. 4) rather than the KS relation, because a time scale can be interpreted physically and because it reduces the dynamic range of the y-axis, which better highlights differences. We call this the t_{gas} relation.

Increasing the metallicity from $\log_{10} Z/Z_{\odot} = -0.3$ (orange) to 0.3 (green) barely impacts the gas consumption time for neutral gas, decreases (increases) that for atomic (molecular) gas, particularly at low surface densities. This mainly reflects the higher molecular fractions for higher metallicities due to the increased abundance of dust grains that catalyse molecule formation and shield dissociating radiation, and the increase in metal cooling that enables more gas to cool to the low temperatures required for molecule formation. Despite these differences, all metallicity variations fall within the range spanned by the observations (which, as discussed above, may however suffer from a metallicity-dependent bias). We now turn to the analysis of the effect of the spatial averaging method.

4.1.3 The spatial averaging scale

Fig. 5 shows the neutral, molecular and atomic t_{gas} relations for the fiducial simulation averaged over different spatial scales. For spatial bins of size $\leq 3 \text{ kpc}$, the neutral and atomic gas consumption times steepen rapidly at low surface densities, indicating very inefficient SF. However, for larger spatial bin sizes this break moves to smaller surface densities, which agrees with observations (e.g. Onodera et al. 2010). For the largest bin size there is no break at all and the gas consumption time is reasonably well approximated by a power law of the surface density. This indicates that the larger spatial bins contain subregions with higher surface density and smaller, local gas consumption times. For a given spatial bin size, the break occurs at smaller molecular than atomic surface densities. As expected, the scatter in $t_{\text{gas},i}$ decreases when averaging over larger spatial scales. At the largest spatially-averaging scales, the $t_{\text{gas,H+H}_2}$ relation is close to the observations of K98 for entire galaxies (dot-dashed lines). The normalisation of the t_{gas} relations are lower for larger spatially-averaging scales, which is in agreement with the different normalisations of the galaxy-averaged KS relation of Kennicutt (1998) and the spatially-resolved KS relation of Kennicutt et al. (2007) who found a difference in normalisation similar to what we find.

4.1.4 Azimuthally-averaged vs. spatially-resolved t_{gas} relations

Fig. 6 shows a comparison between the spatially-resolved (i.e. local) and the azimuthally-averaged t_{gas} relations for bin sizes of 750 pc and for neutral, molecular and atomic hydrogen, respectively. The predicted spatially-resolved relations are compared with spatially-resolved observations (B08; B10) and the predicted azimuthally-averaged relations are compared with azimuthally-averaged observations (Schrubba et al. 2011; hereafter S11). We corrected the surface densities of S11 by removing the factor for helium and heavy elements. For reference, we also show the observed galaxy-averaged $t_{\text{gas},i}$ (K98). The simulations are overall in good agreement with the observations for the different phases and averaging methods. Compared to S11 the predicted gas consumption times are slightly too large at low surface densities, but S11 only use data from the inner

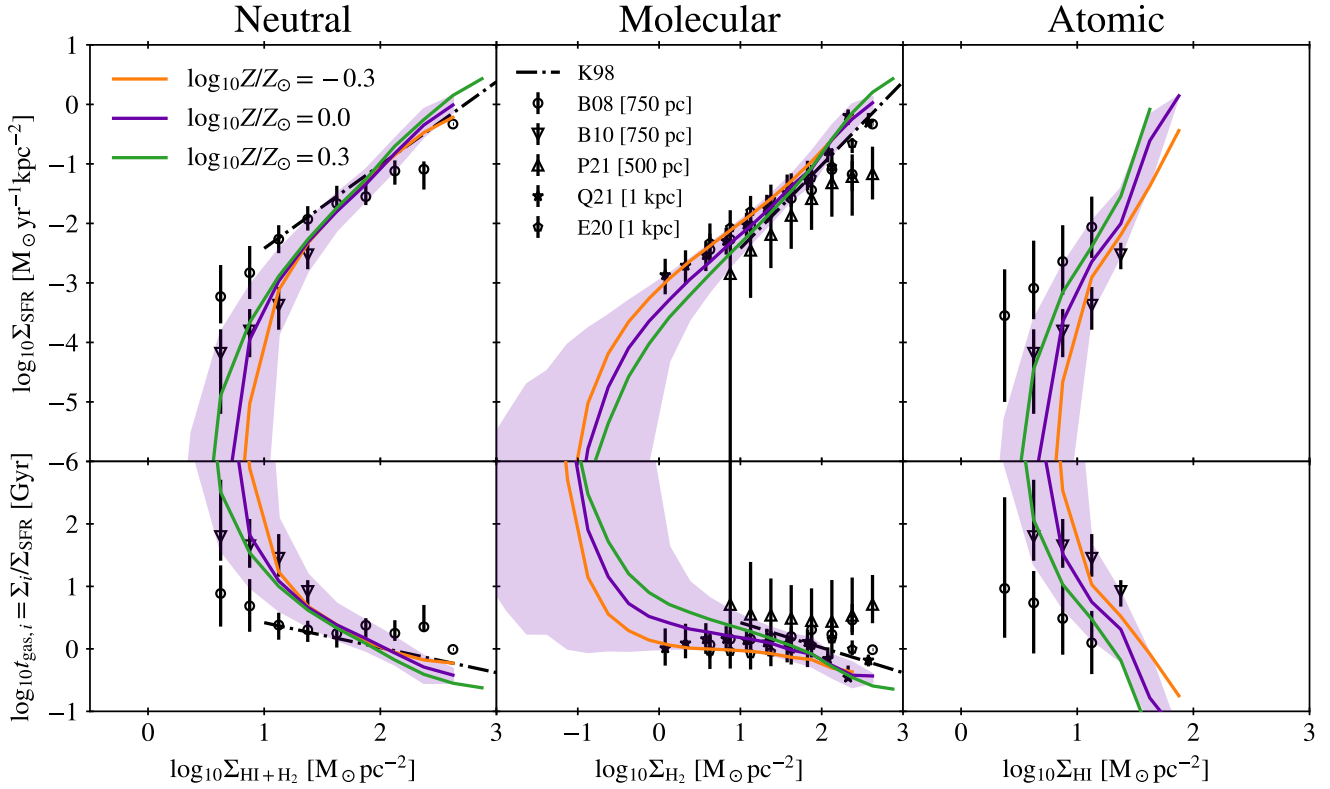


Figure 4. Comparison of the (face-on) spatially-resolved median star formation rate surface density (i.e., the Kennicutt-Schmidt (KS) relation) (top) and the median gas consumption time t_{gas} (bottom) as a function of gas surface density for neutral (left), molecular (middle) and atomic (right) hydrogen. The results are averaged over cells of size 750 pc for different metallicities (different colours). The shaded region shows the 16th and 84th percentile scatter for the $Z = Z_{\odot}$ simulation. We compare with observations of galaxy-averaged KS relations (K98), spatially-resolved atomic and molecular hydrogen surface densities (B08; B10), and spatially-resolved molecular hydrogen surface densities (P21; Q21; E20). The error bars show the 16th and 84th percentile scatter. The simulations are in good agreement with the data. A higher metallicity shifts the cut-off in the neutral and atomic KS relations to lower lower surface densities, while the cut-off in the molecular KS relation shifts to higher surface densities.

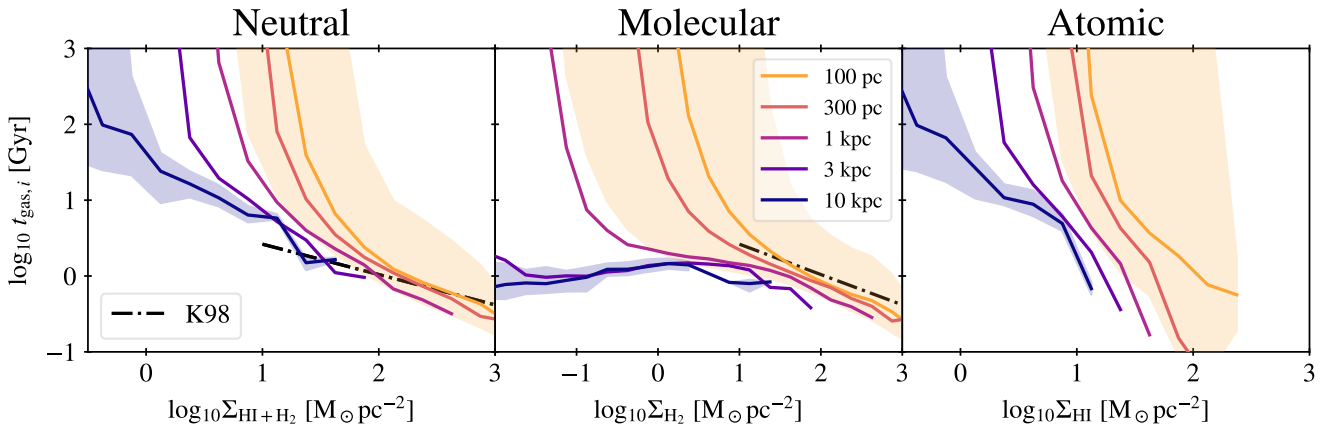


Figure 5. Comparison of the spatially-resolved relation between the median gas consumption time and the gas surface density for neutral (left), molecular (centre) and atomic hydrogen (right), averaged over different spatial scales (different coloured lines) for the fiducial simulation with $Z = Z_{\odot}$. For the largest (10 kpc) and lowest (100 pc) bin sizes the shaded regions indicate the 16th and 84th percentile scatter. At low surface densities the relations steepen. For smaller (larger) observed patches, this break in the t_{gas} relation becomes more (less) pronounced and shifts to higher (lower) surface densities. For the largest spatial patches, there is no break.

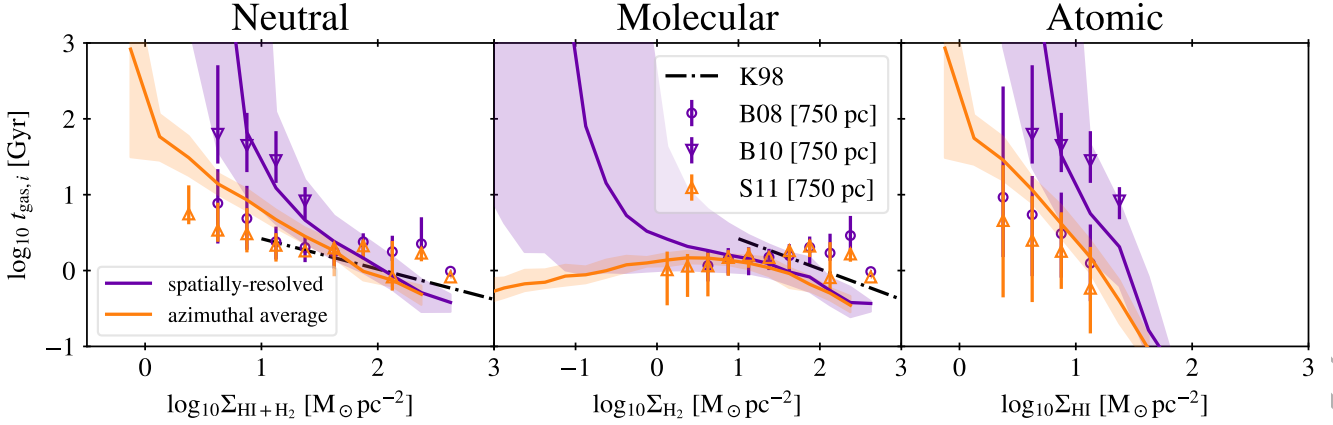


Figure 6. Comparison of the spatially-resolved (purple) and azimuthally-averaged (orange) median t_{gas} relations for all neutral (left), molecular (centre) and atomic hydrogen (right) for the fiducial simulation with solar metallicity. The shaded regions indicate the 16th - 84th percentile scatter. We compare with disk-averaged observations of spiral and star-forming galaxies (dot-dashed line; K98), azimuthally-averaged observations of S11 and the spatially-resolved data from B08 and B10. The error bars show the 16th-84th percentile scatter. The observations and simulations use the same spatial and azimuthal resolution of 750 pc. The spatially-resolved relations are much steeper than the azimuthally-averaged ones, particularly for neutral and atomic gas. The simulations agree with both types of observed relations.

Table 1. Resolution variations

Simulation	particle mass (M_{\odot})	softening (pc)	α_{crit}
fid	10^5	200	1.0
highres8fixedalpha	1.25×10^4	100	1.0
lowres8fixedalpha	8×10^5	400	1.0
lowres64fixedalpha	6.4×10^6	800	1.0
lowres512fixedalpha	5.12×10^7	1600	1.0
highres8scaledalpha	1.25×10^4	100	4.0
lowres8scaledalpha	8×10^5	400	0.25
lowres64scaledalpha	6.4×10^6	800	0.00625
lowres512scaledalpha	5.12×10^7	1600	0.001562

regions of galaxies (i.e. mostly the same data as B08, and partly B10).

While the spatially-resolved and azimuthally-averaged gas consumption time scales converge at high surface densities, they differ strongly at low surface densities because the spatially-resolved t_{gas} relations are steeper. Comparing Figs. 5 and 6, we see that the azimuthally-averaged relations using a radial bin size of 750 pc are similar to the spatially-resolved relations using a bin size of $\sim 3 - 10$ kpc, which is comparable to the radial disk scale length of 4.3 kpc.

For the remainder of the paper, we will focus on the neutral and molecular t_{gas} relations. We now turn to resolution tests.

4.1.5 Resolution tests

Equation (27) indicates that the selection of star-forming gas depends directly on the numerical resolution. This means that for a fixed value of α_{crit} different regions in $(T, n_{\text{H}}, \sigma_{\text{turb}})$ space are selected for different resolutions. For a fixed α_{crit} but higher (lower) resolution, gas is selected with higher (lower) densities, lower (higher) temperatures and lower (higher) turbulent velocity dispersions. We therefore look at the convergence with resolution in two different settings: a fixed α_{crit} and a scaled α_{crit} ($\alpha_{\text{crit}} \propto M_{\text{particle}}^{-2/3}$). For fixed α_{crit} the physical conditions need to be such that the mini-

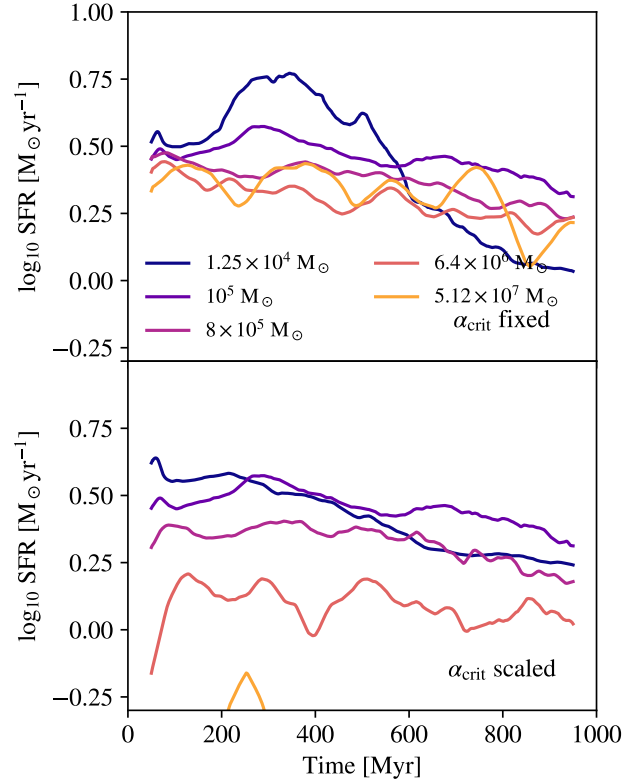


Figure 7. Comparison of the SFHs for simulations with different resolutions (different colours) and either constant α_{crit} (top panel) or scaled α_{crit} ($\alpha_{\text{crit}} \propto M_{\text{particle}}^{-2/3}$; bottom panel). The $M_{\text{particle}} = 10^5 M_{\odot}$ simulation is the same in both panels. The lines show the 100 Myr moving average SFR. For scaled α_{crit} the SFR is systematically lower for larger particle masses, while for a fixed α_{crit} the results are insensitive to the resolution.

mum resolved mass is unstable, while for scaled α_{crit} the same mass is required to be unstable when the resolution changes and hence the same regions in $(T, n_{\text{H}}, \sigma_{\text{turb}})$ space are selected. In Table 1 we list the different simulations used for the resolution tests.

Fig. 7 shows the SFHs for the different resolution variations.

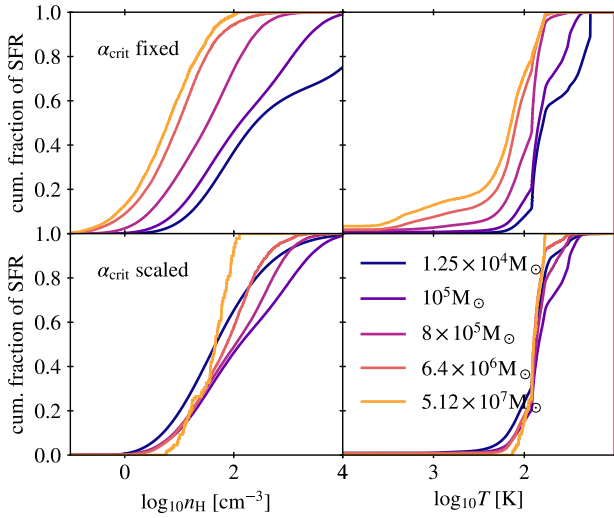


Figure 8. Comparison of the cumulative fraction of the SFR as a function of the gas density (left) and temperature (right) for simulations with different resolutions (different colours) and either constant α_{crit} (top) or scaled α_{crit} ($\alpha_{\text{crit}} \propto M_{\text{particle}}^{-2/3}$; bottom). Because of the explicit resolution dependence of the instability criterion for fixed α_{crit} , the convergence is much better if α_{crit} is scaled with the resolution.

We do not expect exact convergence because the initial density profile is stochastically realized and both SF and stellar feedback are implemented stochastically. The SFHs converge for the simulations with fixed α_{crit} as compared to those with a scaled α_{crit} . For the higher-resolution simulations the differences in the SFHs are within 0.2 dex, similar to the impact of varying the metallicity by a factor of two (see Fig. 3) and they do not vary systematically with the resolution, however, the highest resolution simulation shows that the SFR is boosted until sufficient energy has been released to trigger a strong outflow, after which feedback again becomes efficient and the SFR declines more rapidly. This is an indication that the model is reaching its limit. The convergence is good for all particle masses for fixed α_{crit} and not good for scaled α_{crit} . For scaled α_{crit} , simulations with even lower resolutions predict lower SFRs (the simulation with $5.12 \times 10^7 M_{\odot}$ and scaled α_{crit} does only form a few stars). At low resolutions the scaled α_{crit} simulations have the problem that the density of gas that becomes star-forming (following equation 31) is much higher than $M_{\text{particle}}/\epsilon^3$, the density at which gravitational softening becomes important.

Fig. 8 shows the cumulative fraction of the SFR⁸ as a function of density and temperature. As expected, in simulations using a fixed α_{crit} stars form in gas with higher densities and lower temperatures if the resolution is increased. The simulations with scaled α_{crit} converge better in terms of the densities and temperatures of the gas from which the stars are predicted to form because they select gas with the same physical properties. However, Fig. 7 shows that better convergence of the densities and temperatures of the gas from which stars form does not imply better convergence of the SFR.

The stellar birth densities and temperatures in the simulations are not particularly interesting properties because they are not observable. In reality, stars form at much higher densities, which remain completely unresolved in the simulations. In fact, if the physical properties of the gas particles that are converted into stars were

⁸ Calculated as $\sum_i \text{SFR}(n_{\text{H},i} < n_{\text{H}}) / \text{SFR}$ and $\sum_i \text{SFR}(T_i > T) / \text{SFR}$.

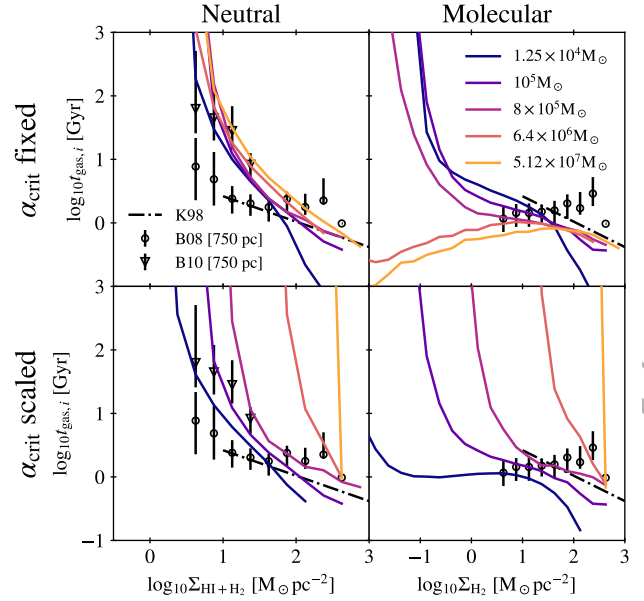


Figure 9. Comparison of the spatially resolved (750 pc) neutral (left) and molecular (right) t_{gas} relations for simulations with different resolutions (different colours) and different instability criteria (top and bottom). For reference, the dot-dashed line shows the galaxy-averaged t_{gas} relation from K98 and the data points show the spatially-resolved t_{gas} relations from B08 and B10. The neutral t_{gas} relation is close to converged with the numerical resolution for fixed α_{crit} , but if α_{crit} is scaled with the resolution, then the relation shifts to longer gas consumption time scales if the resolution is decreased. For the molecular t_{gas} relation the convergence is poor at low surface densities, with higher resolution resulting in longer gas consumption time scales, but at high surface densities ($\Sigma_{\text{H}_2} > 10 M_{\odot} \text{pc}^{-2}$) the results are less sensitive to the resolution.

converged, then increasing the resolution would not result in more realistic properties of star-forming gas and hence would not open new avenues to test the models such as predictions for observational tracers of gas with higher densities. Our goal is therefore not to obtain converged results for the properties of resolution elements that are being converted into stars, but to obtain converged predictions for observables probing fixed physical scales, such as the KS relations.

Fig. 9 shows how the spatially-resolved gas consumption time scales vary with the surface density for the neutral (left) and molecular (right) gas for different numerical resolutions (different colours) and the two α_{crit} variations (different line styles). For fixed α_{crit} the convergence of the neutral t_{gas} is good (within 0.2 dex). The scaled α_{crit} performs less well: the t_{gas} relation shifts to longer consumption time scales if the resolution decreases. At low surface densities the molecular t_{gas} relation shows poor convergence for both fixed α_{crit} and scaled α_{crit} . However, for fixed α_{crit} the convergence is reasonable for $\Sigma_{\text{H}_2} > 10 M_{\odot} \text{pc}^{-2}$.

4.1.6 The t_{gas} relations for different gas phases

Fig. 10 shows the hydrogen surface density fractions of ionized, atomic and molecular versus the neutral hydrogen surface density. Lines and data points indicate, respectively, locally and azimuthally averaged quantities, both using bins of 750 pc at our fiducial resolution. As the neutral surface density increases, the dominant species changes from ionized to atomic to molecular. The transitions are at slightly lower surface densities for the azimuthally averaged case and

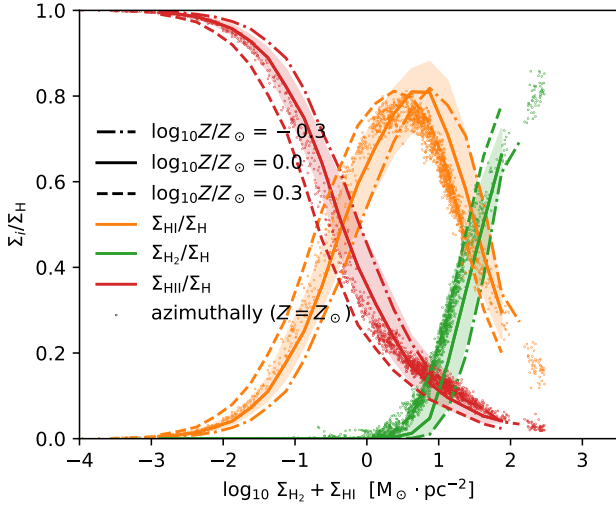


Figure 10. The contributions of different species of hydrogen (different colours) to the total (face-on) hydrogen surface densities as a function of the neutral hydrogen surface density for different metallicities (different line styles). The lines show the medians and the shaded regions show the 16th and 84th percentile scatter for the case of solar metallicity, averaged over 750 pc regions. The individual points show the azimuthally averaged values, using 750 pc wide rings. At the highest surface densities, $\Sigma_{\text{H}_2+\text{H}_I} \sim 10^2 \text{ M}_\odot \text{ pc}^{-2}$, the contribution from ionized hydrogen is still not fully negligible. At low surface densities, $\Sigma_{\text{H}_2+\text{H}_I} \lesssim 1 \text{ M}_\odot \text{ pc}^{-2}$, it is possible to ignore the contribution from molecular hydrogen for spatially resolved and azimuthally averaged measurements. The transitions shift to higher surface densities for lower metallicities.

they shift to lower surface densities if the metallicity is increased. We note that the contribution of ionised hydrogen, which is typically ignored, is significant at all neutral surface densities. Even at our highest densities, $\Sigma_{\text{H}_I+\text{H}_2} \sim 10^2 \text{ M}_\odot \text{ pc}^{-2}$, it still accounts for nearly 10 per cent of the gas mass.

Fig. 11 directly compares the t_{gas} relations for the three different hydrogen species and their sum for the fiducial model, this includes some of the lines from Fig. 6. We show both the spatially resolved (top panel) and the azimuthally averaged t_{gas} relation (bottom panel). The neutral and total hydrogen t_{gas} relations are very close, particularly at high surface densities ($\Sigma_{\text{H}_I+\text{H}_2} \gg 10 \text{ M}_\odot \text{ pc}^{-2}$). The atomic t_{gas} relation lies below both and deviates particularly strongly for $\Sigma_{\text{H}_I} \gtrsim 10^{1.5} \text{ M}_\odot \text{ pc}^{-2}$. The molecular relation deviates the most, giving much shorter t_{gas} for surface densities $\Sigma_{\text{H}_2} \lesssim 10^{1.5} \text{ M}_\odot \text{ pc}^{-2}$. While for molecular gas the consumption time scale is about 1 Gyr across the full range of surface densities, for the other species the gas consumption time scale increases rapidly with decreasing surface density and this trend is stronger for spatially-resolved than for azimuthally-averaged measurements.

4.1.7 The impact of gas metallicity

Fig. 1 showed that at a fixed density, lower metallicity gas has higher equilibrium temperatures. Therefore, for a fixed turbulent velocity dispersion the gravitational instability criterion applied on a fixed mass scale (i.e., for a fixed numerical resolution), will tend to select gas with higher densities and higher temperatures if the metallicity is lower. This can be seen from Fig. 1 by comparing the intersections of an instability line with the thermal equilibrium lines for different metallicities. Motivated by this observation, in Fig. 12

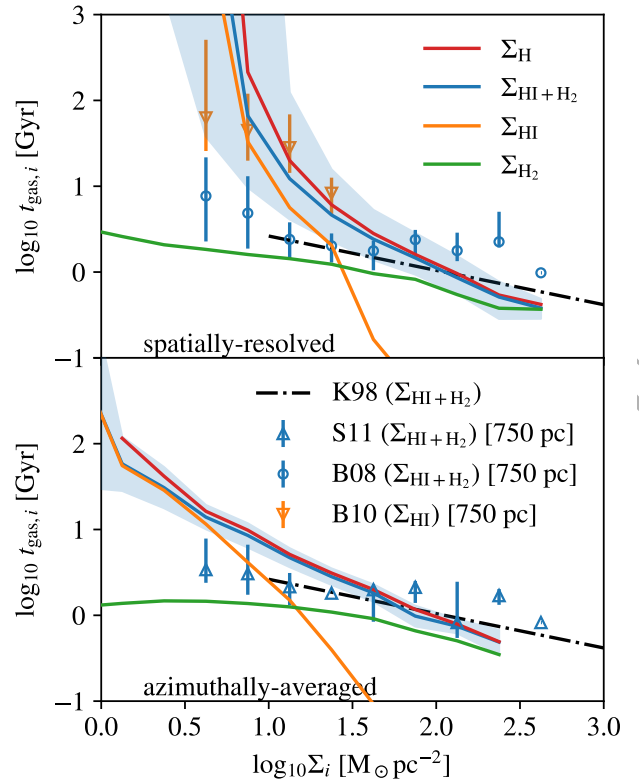


Figure 11. Comparison of the median t_{gas} relations for different gas phases (different colours) at scales of 750 pc for the fiducial model (note that many of the lines were already shown in Fig. 6). The shaded regions show the 1 sigma scatter for the neutral gas relation. The top and bottom panels show, respectively, spatially resolved and azimuthally averaged measurements. The simulations are compared with observations of disk averaged relations in spiral and star-forming galaxies (dot-dashed line; K98) and with spatially-resolved observations for neutral gas and atomic gas (top; data points with error bars indicating 16th and 84th percentile scatter; B08;B10) and azimuthally-averaged neutral gas observations (bottom; error bars indicating 16th and 84th percentile scatter; S11). While the molecular gas consumption time scale is nearly constant at ≈ 1 Gyr, for atomic, neutral and total hydrogen t_{gas} is typically larger and increases rapidly towards lower surface densities, particularly for the case of spatially resolved measurements.

we show the cumulative contributions of different densities (left) and temperatures (right) to the total SFR for different metallicities. For high metallicities ($Z \geq 0.1Z_\odot$) the SFR is dominated by cold gas ($T \lesssim 10^2 \text{ K}$), while for very low metallicities ($Z \leq 10^{-2}Z_\odot$) the SFR is dominated by warm ($T \gtrsim 10^3 \text{ K}$) gas. We emphasize that while the trend with metallicity is generic, the fractions of stars forming at different densities and temperatures depend explicitly on the numerical resolution, because we evaluate the gravitational instability criterion at the minimum resolved mass. An increase in numerical resolution leads to a smaller mass that has to become unstable, which requires higher densities and/or lower temperatures (see Fig. 1).

Fig. 13 shows the metallicity dependence of the spatially-resolved neutral (left) and molecular (right) KS relations. Unlike Fig. 12, these predictions are for observables and do not depend explicitly on the numerical resolution, because we average over a fixed spatial scale (750 pc) that is not explicitly tied to the resolution (which does not automatically mean the results are converged with the resolution, but they at least could be). Decreasing the metallicity from solar to primordial shifts the break in the neutral KS relation

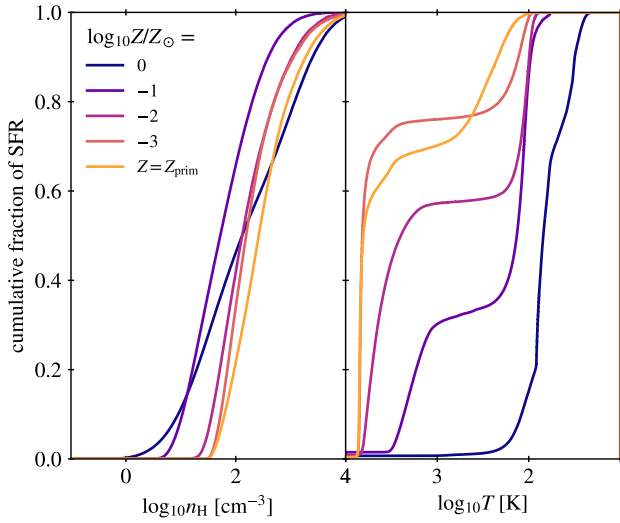


Figure 12. Comparison of the cumulative contribution of gas with different hydrogen number densities (left) and temperatures (right) to the total SFR for simulations assuming different metallicities (different colours). For lower metallicities star formation occurs in gas with higher temperatures and lower densities. This can be understood by comparing a gravitational instability line with the thermal equilibrium lines for different metallicities in the temperature-density diagram shown in Fig. 1.

to around one order of magnitude higher neutral surface densities, but has little effect at higher surface densities (in agreement with simulations with a different SF criterion, e.g. Richings & Schaye 2016). The molecular KS relation is affected more strongly, it shifts to lower surface densities and has a much higher Σ_{SFR} at fixed Σ_{H_2} . We caution, however, that at low surface densities the molecular KS law is sensitive to the resolution (see Fig. 9). The difference in the molecular KS relation is dominated by the metallicity dependence of the molecular fraction, at lower abundances and a constant density there is less dust to catalyse molecule formation and shield dissociating radiation. In Appendix A, we show the neutral and molecular KS relations for the temperature ceiling SF criterion. Simulations using that criterion give a similar dependence on the metallicity although the break in the neutral KS relation is more sensitive to the metallicity than if the gravitational instability criterion is used. This is expected because while for the fiducial criterion a lower metallicity may result in the selection of higher density but warmer gas (see Fig. 1), for the temperature ceiling criterion the gas must be cold and in thermal equilibrium, meaning that its density must increase further. Overall, these results imply that the use of an H_2 -dependent SF threshold can prevent the formation of stars in gravitationally unstable gas at low metallicities. This is due to the fact that H_2 is only a good tracer of SF at higher metallicities and is not the cause of SF. The right panel of Fig. 13 suggests that simulations that restrict star formation to molecular gas, would need to use a SF efficiency that decreases strongly towards low metallicities in order to be consistent with the molecular KS law. This leads us to conclude that it is not advisable to use an H_2 -dependent SF threshold in simulations of galaxy formation. In Appendix B, Fig. B3 shows that the SFH is lower for lower metallicities but for metallicities $\log_{10} Z/Z_{\odot} \leq -2$ the SFH does not change significantly.

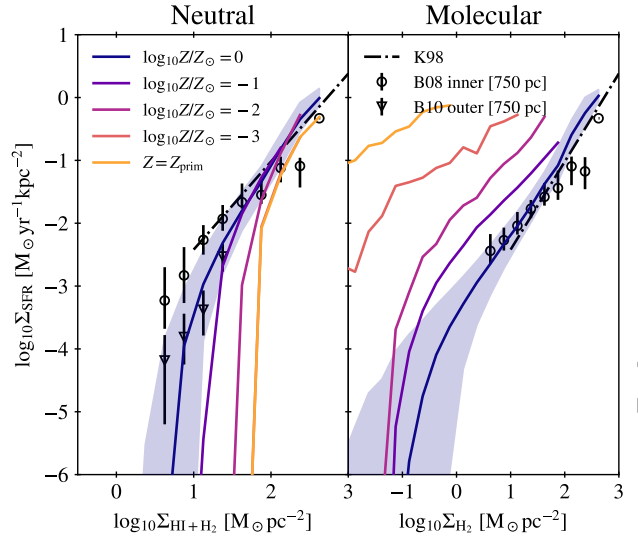


Figure 13. Comparison of the neutral (left) and molecular (right) spatially-resolved (750 pc) median KS relations for simulations assuming different metallicities (different colours). The shaded regions show the 1 sigma scatter for the case of solar metallicity. For reference, the disk-averaged t_{gas} relation from K98 (dot-dashed line) and the spatially-resolved t_{gas} relations from B08 and B10 are shown. The cut-off in the neutral KS relation shifts to about 1 dex higher surface density going from solar metallicity to primordial abundances, but the relations converge at high surface densities. The molecular KS relation changes dramatically, with lower metallicities giving higher Σ_{SFR} at a fixed molecular surface density.

4.2 The impact of subgrid physics variations

In this section, we focus on the impact of changing subgrid physics parameters on the spatially-resolved t_{gas} relations. In particular, we will compare variations in the SF criterion (§4.2.1), the SF efficiency (§4.2.2), and the SN feedback energy (§4.2.3). Table 2 lists all the different simulation variations.

4.2.1 The star formation criterion

Here we show the impact of the different SF criteria discussed in Section 3 on the t_{gas} relations. We compare different normalisations of the gravitational instability criterion, different temperature ceilings and different density thresholds.

Fig. 14 shows the SFHs for the simulations using different SF criteria. The SFHs have been averaged over 100 Myr to remove the short-time scale scatter. All simulations use our fiducial resolution ($M_{\text{particle}} = 10^5 M_{\odot}$) and solar metallicity. Despite the huge differences in SF criteria, the SFHs are remarkably similar. Only the most restrictive criteria, i.e. $\alpha_{\text{crit}} \leq 0.1$ and $n_{\text{H,crit}} \geq 100 \text{ cm}^{-3}$ result in significantly different SFHs. These simulations predict lower SFRs, probably because the numerical resolution becomes a bottleneck for reaching the densities required for SF. Once SF starts, the densities will be too high for feedback to be effective, causing a larger amount of SF that only stops when the densest gas has been consumed and/or sufficient SN energy has been released to overcome the large radiative losses. This then causes a more rapid decline in the SFH (similar to the highest resolution simulation, see Fig. 7). Relaxing the criteria a lot does not produce a significant change in the SFH, which indicates that the SFH is determined mostly by self-regulation through stellar feedback.

Fig. 15 shows the spatially-resolved neutral (top) and molecular (bottom) t_{gas} relations for the different SF criteria. As expected, the

Table 2. Variations in the subgrid parameters studied in this work. The columns list, from left to right, the simulation identifier, the adopted SF criterion (SF crit.), the energy per core collapse SN (E_{CCSN}), the SF efficiency per free fall time in the Schmidt law (ϵ). Parameter values that differ from the fiducial model are shown in bold face. All simulations use our fiducial particle mass ($10^5 M_{\odot}$), solar metallicity and a kinetic feedback fraction of $f_{\text{kin}} = 0.1$.

Simulation	SF crit.	E_{CCSN} (10^{51} erg)	ϵ
fid	$\alpha_{\text{crit}} = 1$	2.0	0.01
alphacrit01	$\alpha_{\text{crit}} = \mathbf{0.1}$	2.0	0.01
alphacrit03	$\alpha_{\text{crit}} = \mathbf{0.3}$	2.0	0.01
alphacrit3	$\alpha_{\text{crit}} = \mathbf{3}$	2.0	0.01
alphacrit10	$\alpha_{\text{crit}} = \mathbf{10}$	2.0	0.01
alphacrit30	$\alpha_{\text{crit}} = \mathbf{30}$	2.0	0.01
T1e4K	$T_{\text{crit}} = \mathbf{10^4 K}$	2.0	0.01
T3e3K	$T_{\text{crit}} = \mathbf{3 \times 10^3 K}$	2.0	0.01
T1e3K	$T_{\text{crit}} = \mathbf{10^3 K}$	2.0	0.01
T3e2K	$T_{\text{crit}} = \mathbf{3 \times 10^2 K}$	2.0	0.01
T1e2K	$T_{\text{crit}} = \mathbf{10^2 K}$	2.0	0.01
nH1e-2cm-3	$n_{\text{H,crit}} = \mathbf{0.01 cm^{-3}}$	2.0	0.01
nH1e-1cm-3	$n_{\text{H,crit}} = \mathbf{0.1 cm^{-3}}$	2.0	0.01
nH1e0cm-3	$n_{\text{H,crit}} = \mathbf{1 cm^{-3}}$	2.0	0.01
nH1e1cm-3	$n_{\text{H,crit}} = \mathbf{10 cm^{-3}}$	2.0	0.01
nH1e2cm-3	$n_{\text{H,crit}} = \mathbf{100 cm^{-3}}$	2.0	0.01
nH1e3cm-3	$n_{\text{H,crit}} = \mathbf{1000 cm^{-3}}$	2.0	0.01
SFE0.001	$\alpha_{\text{crit}} = 1$	2.0	0.001
SFE0.1	$\alpha_{\text{crit}} = 1$	2.0	0.1
SFE1.0	$\alpha_{\text{crit}} = 1$	2.0	1.0
ECCSN0.25e51	$\alpha_{\text{crit}} = 1$	0.25	0.01
ECCSN0.5e51	$\alpha_{\text{crit}} = 1$	0.5	0.01
ECCSN1e51	$\alpha_{\text{crit}} = 1$	1	0.01
ECCSN4e51	$\alpha_{\text{crit}} = 1$	4	0.01
ECCSN8e51	$\alpha_{\text{crit}} = 1$	8	0.01

results diverge towards low surface densities. There are nevertheless large differences between the models in the surface density range for which there are observational constraints. The SF criterion mainly affects the surface density below which t_{gas} becomes very large. If the SF criterion is too strict, then t_{gas} already diverges from the data at surface density for which there are still good constraints. If the criterion is too relaxed, then the neutral gas KS law is too shallow compared with the B10 data for outer disks. The most relaxed and the most strict gravitational instability criteria predict, respectively, too short and too long gas consumption time scales, particularly for molecular gas. Only the fiducial $\alpha_{\text{crit}} = 1$ and the $\alpha_{\text{crit}} = 3$ models appear consistent with the data. Only the lowest temperature ceiling ($T_{\text{crit}} = 10^2$ K) agrees with the observed molecular t_{gas} and the data rule out all density threshold criteria. Many of the simulations predict t_{gas} relations with the wrong shape, which suggests that for most SF criteria tuning the SFE parameter would not resolve the discrepancy with the data (for $T_{\text{crit}} = 3 \times 10^2$ K or 10^3 K the SFE parameter could be tuned).

Overall, we find that in order to choose between subgrid SF criteria we should not focus on the SFH as this differs very little for vastly different criteria. This implies that the galaxy-averaged KS relation is also not suitable, at least not for idealized, isolated galaxies. Furthermore, we should not choose based on properties of the gas evaluated at the resolution scale, like stellar birth densities and temperatures, because these are typically not converged and are not measurable. Instead, we find that a suitable way to determine

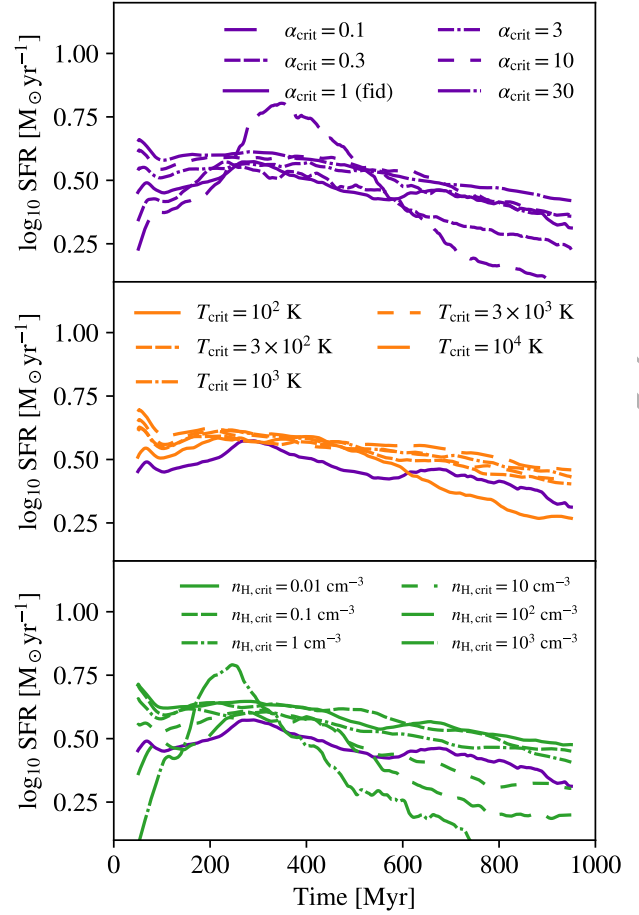


Figure 14. Comparison of the SFHs averaged over 100 Myr time scales for simulations using different SF criteria: a gravitational instability criterion (purple), a temperature ceiling (orange) or a density threshold (green). Different line styles correspond to different parameter values. The fiducial model is the solid purple curve, which is repeated in every panel. The SFHs are remarkably similar for most simulations. Only the three simulations with very strict SF criteria, i.e. $\alpha_{\text{crit}} = 0.1$, $n_{\text{H,crit}} = 10^2 \text{ cm}^{-3}$, and $n_{\text{H,crit}} = 10^3 \text{ cm}^{-3}$, predict a significantly different SFH.

whether SF criteria are appropriate is to compare the spatially-resolved t_{gas} (or KS) relations.

4.2.2 The star formation efficiency

The SFH averaged on 100 Myr time scales is not strongly affected by the SFE, except for the initial ~ 100 Myr (see Appendix B, Fig. B1). Because the total SFR is relatively unaffected by the choice of SFE, it is not a suitable observable for the calibration of ϵ . As was the case for the SF criterion, the KS relations, particularly the spatially-resolved, molecular ones, are however sensitive to the value of ϵ .

Fig. 16 shows the impact that the SFE, ϵ , has on the neutral (left) and molecular (right) t_{gas} relations. Besides the fiducial simulation with $\epsilon = 0.01$, we show simulations with $\epsilon = 0.001, 0.1$ and 1 . For high SFEs the gas consumption time scales increase very steeply at low surface densities, creating a break in the t_{gas} relations. At higher surface densities the trend is as expected: a higher SFE results in a shorter gas consumption time scale. However, at fixed surface density the increase of t_{gas} with ϵ is sublinear. Increasing the SFE by a factor of 1000 only results in a decrease of $t_{\text{gas,H}_2}$ by a factor ~ 100 , and for neutral gas the change in the normalisation

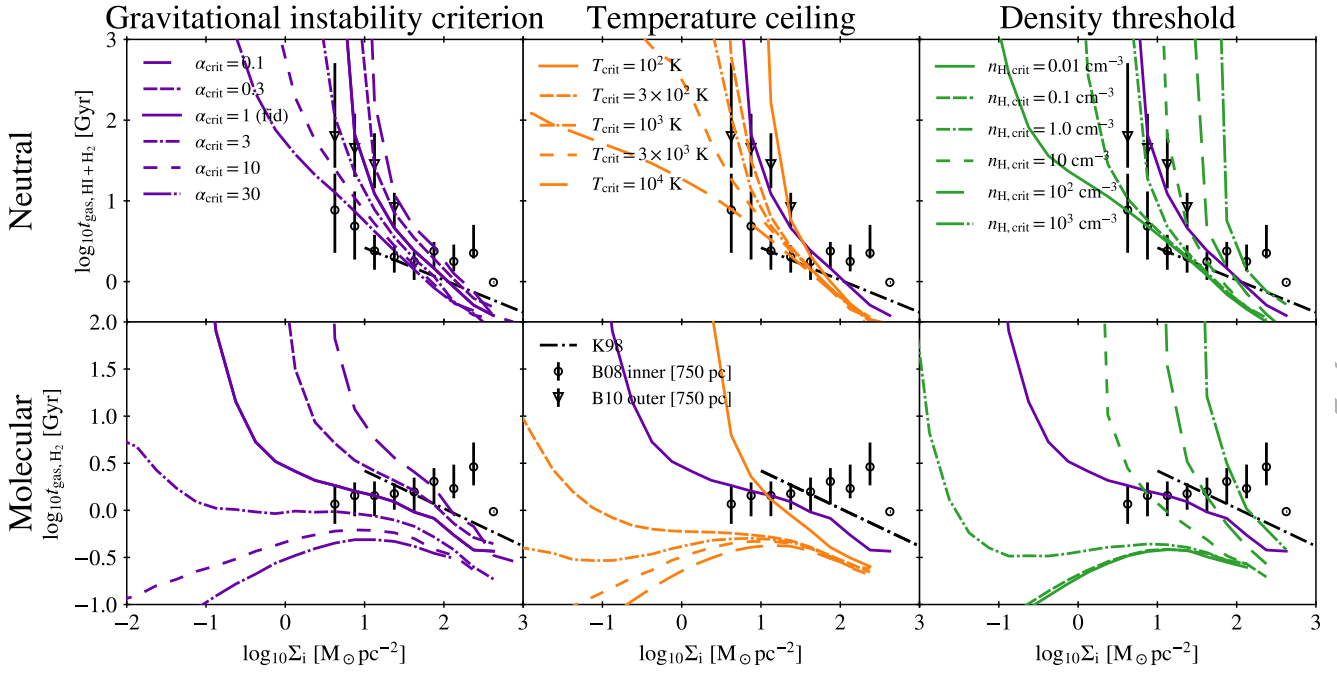


Figure 15. Comparison of the spatially resolved (750 pc) neutral (top) and molecular (bottom) median t_{gas} relations for different SF criteria. The left, middle, and right panels show parameter variations for, respectively, the gravitational instability criterion, a temperature ceiling, and a density threshold. The solid purple line indicates the fiducial model, $\alpha_{\text{crit}} < 1$, which is repeated in every panel. Comparison with the data rules out gravitational instability criteria with $\alpha_{\text{crit}} \ll 1$ or $\alpha_{\text{crit}} \gg 1$ and also all the models with a temperature ceiling or density threshold.

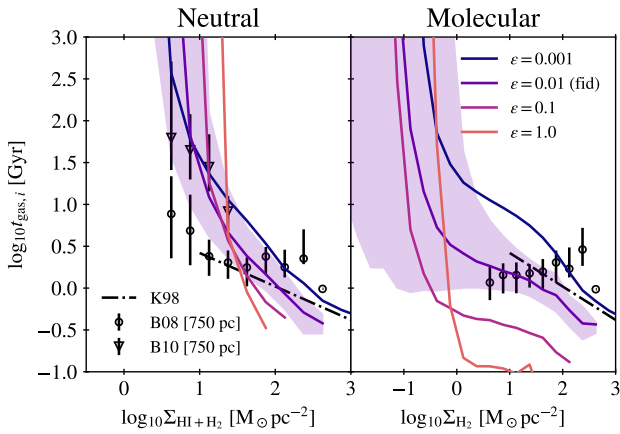


Figure 16. Comparison of the spatially-resolved (750 pc) neutral (left) and molecular (right) median t_{gas} relations for different star formation efficiencies (SFEs) per free fall time, ϵ (different colours). The shaded regions show the 16th and 84th percentile scatter for the fiducial value of $\epsilon = 0.01$. For reference, the disk-averaged t_{gas} relation from K98 (dot-dashed line) and the spatially-resolved t_{gas} relations from B08 and B10 are shown. At high surface densities, where $t_{\text{gas}} \ll 10$ Gyr, the gas consumption time scale decreases with the SFE, but the dependence is sublinear, particularly for neutral gas.

of the t_{gas} relation is even smaller. The break in the neutral t_{gas} relation shifts to higher surface densities when the SFE is increased. The change in the break of the neutral t_{gas} relation is caused by the slightly increased velocity dispersion and temperature in the simulations with higher SFE, likely because SN feedback is more efficient when stars form more rapidly and at lower densities.

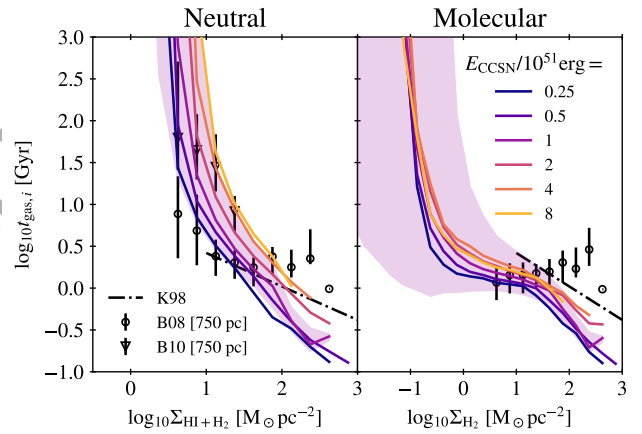


Figure 17. Comparison of the spatially resolved (750 pc) neutral (left) and molecular (right) median t_{gas} relations for different energies per core collapse supernova (different colours). The shaded regions show the 16th and 84th percentile scatter for the fiducial value of $E_{\text{CCSN}} = 2.5 \times 10^{51}$ erg. For reference, the disk-averaged t_{gas} relation from K98 (dot-dashed line) and the spatially-resolved t_{gas} relations from B08 and B10 are shown. The gas consumption time scale increases with the feedback energy, but the dependence is weak except at very low surface densities. At high surface densities the t_{gas} relations, particularly the molecular one, converge with increasing E_{CCSN} .

4.2.3 Supernova feedback

Fig. 17 shows the neutral (left) and molecular (right) spatially-resolved t_{gas} relation for simulations with different values of E_{CCSN} . The neutral t_{gas} relation shifts to about half a dex higher surface densities when E_{CCSN} is increased by a factor of 32 (≈ 1.5 dex). Because the relation steepens towards low surface densities, a con-

stant shift implies that the increase in the neutral gas consumption is smaller at higher surface densities. At $\Sigma_{\text{H}+\text{H}_2} \sim 10^2 \text{ M}_\odot \text{ pc}^{-2}$ t_{gas} increases by about 0.5 dex for an 1.5 dex increase in E_{CCSN} . The molecular gas consumption time scale also increases with E_{CCSN} , and more so at lower surface densities, but the dependence is weaker than for the neutral gas. The two highest feedback energies ($E_{\text{CCSN}}/10^{51} \text{ erg} = 4$ and 8) give nearly identical results. Values of $E_{\text{CCSN}} \sim 1 - 2 \times 10^{51} \text{ erg}$ are most consistent with the data.

The galaxy SF histories (Appendix B, Fig. B3) are much more sensitive to the feedback energy than to the SF parameters. As we have shown, the opposite is the case for the spatially resolved t_{gas} (or KS) relations. The latter are therefore a better calibration target for the SF model, while the SFRs (or, for cosmological initial conditions, the stellar masses) can be used to calibrate the CC SN feedback parameters. However, the t_{gas} relations are not independent of the feedback energy. Furthermore, other subgrid feedback parameters, like the fraction of energy injected in kinetic form, also influence the spatially-resolved KS laws (Chaikin et al. 2023). This means that care needs to be taken when calibrating subgrid models and that it may be necessary to recalibrate the SF model if the feedback prescription is modified.

5 DISCUSSION

5.1 The impact of gravitational softening

To avoid unrealistic two-body scattering, the gravitational force between particles in our simulations is reduced for separations smaller than a fixed gravitational softening length, ϵ . This means that above a certain density, where the separation between particles is smaller than the softening length, we do not correctly simulate the collapse of clouds. The density above which we begin to underestimate the self-gravity of a gas cloud is

$$n_{\text{H,soft}} = \frac{4\pi}{3} \frac{M_{\text{particle}}}{\epsilon^3} \frac{X}{m_{\text{H}}}, \quad (35)$$

$$\approx 16 \text{ cm}^{-3} \left(\frac{M_{\text{particle}}}{10^5 \text{ M}_\odot} \right) \left(\frac{\epsilon}{100 \text{ pc}} \right)^{-3}, \quad (36)$$

where we use $X = 0.756$. For clouds with higher densities self-gravity is not Newtonian (i.e. $F \propto r^{-2}$) while the gravitational instability criterion presented in §3.3 is based on the Newtonian Jeans (1902) mass. In Ploekinger et al. (in prep) we introduced the softened Jeans criteria which extends the Jeans (1902) to be applicable in the case of softened gravitational forces. In Ploekinger et al. (in prep) we studied gravitational (in)stabilities at the resolution limit of Lagrangian simulations assuming only thermal support (i.e. pressure). Here we extend this analysis by including turbulent support and applying the softened Jeans (1902) criterion to the SF criterion.

To investigate the direct impact of softening on the SF criterion, we modify the derivation of the gravitational instability criterion from §3.3 such that it is valid for the case of softened gravity in the limit $\epsilon \gg h$. While the sound and turbulence crossing times across a cloud with radius h and mass m remain identical, the free-fall time of the collapsing cloud in Plummer-softened gravity (i.e.

$\Phi \propto 1/\sqrt{r^2 + \epsilon^2}$) becomes

$$t_{\text{ff,soft}} = -\frac{1}{\sqrt{2Gm}} \int_h^0 \left(\frac{1}{\sqrt{r^2 + \epsilon^2}} - \frac{1}{\sqrt{h^2 + \epsilon^2}} \right)^{-1/2} dr, \quad (37)$$

$$\approx t_{\text{ff}} \sqrt{1 + 2 \left(\frac{\epsilon}{h} \right)^3}, \quad (38)$$

which is the analog of equation (21). In the limit of $\epsilon \gg h$, we obtain

$$t_{\text{ff,soft}} \approx \sqrt{2} t_{\text{ff}} \left(\frac{\epsilon}{h} \right)^{3/2}, \quad (39)$$

$$\approx 64 \text{ Myr} \left(\frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right)^{-1/2} \left(\frac{\epsilon}{100 \text{ pc}} \right)^{3/2} \left(\frac{h}{100 \text{ pc}} \right)^{-3/2}. \quad (40)$$

Therefore, the gas cloud is gravitationally unstable when

$$\sqrt{2} t_{\text{ff}} \left(\frac{\epsilon}{h} \right)^{3/2} < \frac{h}{\sqrt{\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2}}. \quad (41)$$

Setting $h = \lambda_{\text{J,soft}}$ gives the Jeans (1902) length in the limit $\epsilon \gg h$,

$$\lambda_{\text{J,soft}} = \left(\frac{6\pi}{32G\rho} \right)^{1/5} \epsilon^{3/5} (\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2)^{1/5}, \quad (42)$$

$$\approx 3.2 \times 10^2 \text{ pc} \left(\frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right)^{-1/5} \left(\frac{\epsilon}{200 \text{ pc}} \right)^{3/5} \left(\frac{\sqrt{\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2}}{10 \text{ km s}^{-1}} \right)^{2/5}, \quad (43)$$

with a corresponding Jeans (1902) mass of

$$M_{\text{J,soft}} = \frac{4\pi}{3} \rho^{2/5} \left(\frac{6\pi}{32G} \right)^{3/5} \epsilon^{9/5} (\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2)^{3/5}, \quad (44)$$

$$\approx 4.5 \times 10^6 \text{ M}_\odot \left(\frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right)^{2/5} \left(\frac{\epsilon}{200 \text{ pc}} \right)^{9/5} \left(\frac{\sqrt{\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2}}{10 \text{ km s}^{-1}} \right)^{6/5}. \quad (45)$$

This implies that, at densities where self-gravity is strongly softened, the Jeans (1902) mass increases with the density and hence gas that is classified as unstable by the unsoftened instability criterion (discussed in §3.3) may actually be stable. A gas cloud is gravitationally unstable when $m > M_{\text{J}}$, where $m = \langle N_{\text{ngb}} \rangle M_{\text{particle}}$ is the mean mass in the kernel. Therefore, the analogue of equation (27) in the limit $h \ll \epsilon$ becomes

$$\alpha_{\text{soft}} \equiv \frac{\rho^{2/5} \epsilon^{9/5} (\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2)^{3/5}}{G^{3/5} \langle N_{\text{ngb}} \rangle M_{\text{particle}}} < \alpha_{\text{crit,soft}}, \quad (46)$$

$$\alpha_{\text{soft}} \equiv \frac{\rho^{2/5} (\sigma_{\text{th}}^2 + \sigma_{\text{turb}}^2)^{3/5}}{G^{3/5} \langle N_{\text{ngb}} \rangle M_{\text{particle}}^{2/5}} \left(\frac{\epsilon^3}{M_{\text{particle}}} \right)^{3/5} < \alpha_{\text{crit,soft}}, \quad (47)$$

where $\alpha_{\text{crit,soft}}$ is a constant of order unity given by $\alpha_{\text{crit,soft}} = 3(32)^{3/5} / (4\pi(6\pi)^{3/5}) \approx 0.33$ (for unsoftened gravity we had $\alpha_{\text{crit}} \approx 1.3$). In the limiting case of $\sigma_{\text{th}} \gg \sigma_{\text{turb}}$ the instability criterion becomes an upper limit on the temperature,

$$T < 9.7 \times 10^3 \text{ K} \left(\frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right)^{-2/3} \left(\frac{\langle N_{\text{ngb}} \rangle}{65} \right)^{5/3} \times \quad (48)$$

$$\left(\frac{M_{\text{particle}}}{10^5 \text{ M}_\odot} \right)^{5/3} \alpha_{\text{crit,soft}}^{5/3} \left(\frac{\epsilon}{200 \text{ pc}} \right)^{-3}, \quad (49)$$

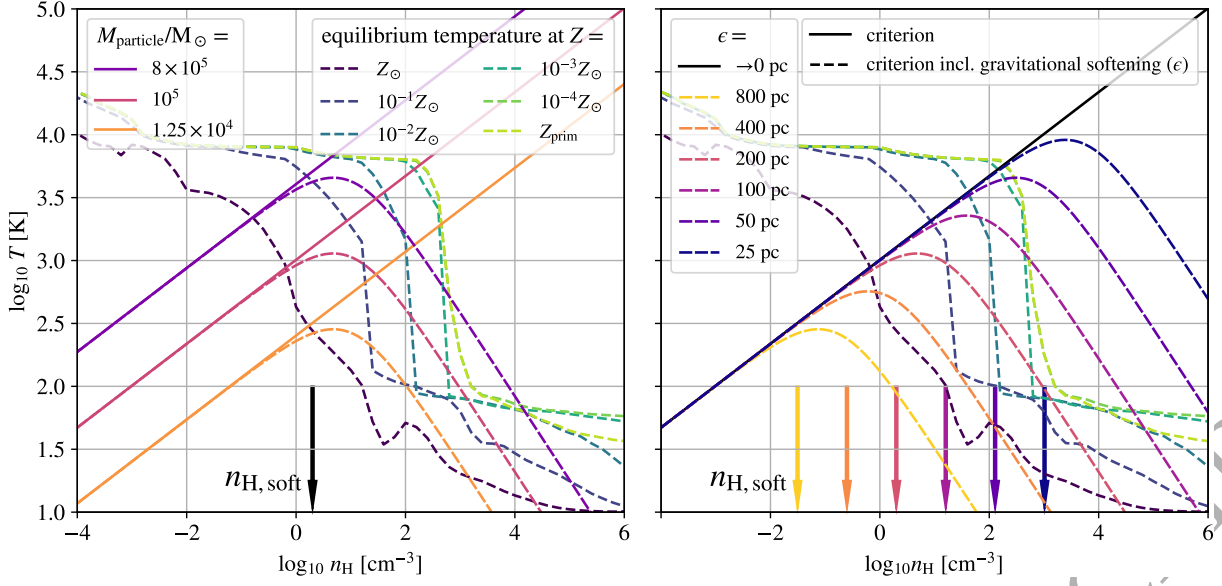


Figure 18. The gravitational instability criterion ($\alpha_{\text{crit}} = 1$) in temperature-density space for softened (dashed lines) and uns softened (solid lines) gravitational forces assuming $\sigma_{\text{th}} \gg \sigma_{\text{turb}}$. The left panel compares different particle masses (different colours) while keeping the ratio $M_{\text{particle}}/\epsilon^3$ fixed and $\alpha_{\text{crit}} = 1$. The right panel compares different gravitational softening lengths (different colours) at our fiducial particle mass of $10^5 M_{\odot}$ and $\alpha_{\text{crit}} = 1$. For reference, we plot the thermal equilibrium temperature curves for different metallicities (dotted lines, different colours) for the Ploekinger et al. (in prep.) cooling tables. The arrows indicate $n_{\text{H,soft}}$, the density above which self-gravity begins to be softened for our fixed ratio of $M_{\text{particle}}/\epsilon^3$ (left) and for different gravitational softening lengths (right). Gas clouds with the minimum resolved mass, $\langle N_{\text{ngb}} \rangle M_{\text{particle}}$, with physical conditions falling in the regions below a given line are gravitationally unstable. In between the solid and dashed lines, such gas clouds should be unstable, but softening keeps them stable. A smaller gravitational softening length moves the softened gravitational instability curve to higher densities.

or

$$T < 9.7 \times 10^3 \text{ K} \left(\frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right)^{-2/3} \left(\frac{\langle N_{\text{ngb}} \rangle}{65} \right)^{5/3} \times \left(\frac{M_{\text{particle}}}{10^5 M_{\odot}} \right)^{2/3} \alpha_{\text{crit,soft}}^{5/3} \left(\frac{M_{\text{particle}}/\epsilon^3}{10^5 M_{\odot}/(200 \text{ pc})^3} \right), \quad (50)$$

where we assumed a mean particle mass $\mu = 1.3$ and a hydrogen mass fraction $X = 0.756$. Note that instability requires a lower temperature if the softening length is larger.

The limiting case of $\sigma_{\text{turb}} \gg \sigma_{\text{th}}$ gives an upper limit on the density of

$$n_{\text{H}} < 7.4 \times 10^3 \text{ cm}^{-3} \left(\frac{\sigma_{\text{turb}}}{5 \text{ km s}^{-1}} \right)^{-3} \left(\frac{\langle N_{\text{ngb}} \rangle}{65} \right)^{5/2} \times \left(\frac{M_{\text{particle}}}{10^5 M_{\odot}} \right)^{5/2} \alpha_{\text{crit,soft}}^{5/2} \left(\frac{\epsilon}{100 \text{ pc}} \right)^{-9/2}, \quad (51)$$

$$\left(\frac{M_{\text{particle}}}{10^5 M_{\odot}} \right)^{5/2} \alpha_{\text{crit,soft}}^{5/2} \left(\frac{\epsilon}{100 \text{ pc}} \right)^{-9/2}, \quad (52)$$

or

$$n_{\text{H}} < 7.4 \times 10^3 \text{ cm}^{-3} \left(\frac{\sigma_{\text{turb}}}{5 \text{ km s}^{-1}} \right)^{-3} \left(\frac{\langle N_{\text{ngb}} \rangle}{65} \right)^{5/2} \times \left(\frac{M_{\text{particle}}}{10^5 M_{\odot}} \right)^{5/2} \alpha_{\text{crit,soft}}^{5/2} \left(\frac{M_{\text{particle}}/\epsilon^3}{10^5 M_{\odot}/(200 \text{ pc})^3} \right)^{3/2}. \quad (53)$$

This means that to resolve instabilities at higher densities the softening length is required to be smaller.

The left panel of Fig. 18 shows the gravitational instability criterion derived in §3.3 in temperature-density space (solid lines) for different values of the particle mass M_{particle} (different colours) and $\alpha_{\text{crit}} = 1$. This is compared with the softened gravitational

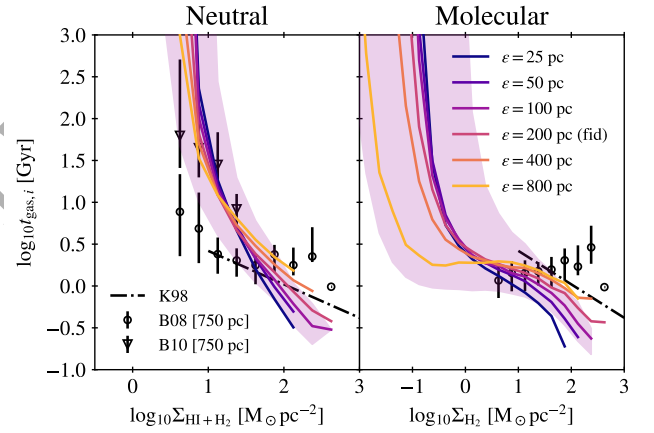


Figure 19. Comparison of the spatially resolved (750 pc) neutral (left) and molecular (right) median t_{gas} relations for different gravitational softening lengths (different colours). The shaded regions show the 16th and 84th percentile scatter for the fiducial value $\epsilon = 100$ pc. For reference, the disk-averaged t_{gas} relation from K98 (dot-dashed line) and the spatially-resolved t_{gas} relations from B08 and B10 are shown. At low surface densities the gas consumption time scale is almost independent of the gravitational softening length as long as it is not too large ($\epsilon \lesssim 400$ pc). At high surface densities the gas consumption time scale decreases systematically with increasing force resolution, though the dependence is weak.

instability criterion in the limit $\sigma_{\text{th}} \gg \sigma_{\text{turb}}$ and assuming $\epsilon \propto M_{\text{particle}}^{1/3}$ (dashed lines). For the softened gravitational instability criterion we derived the limit of $\epsilon \gg h$ and interpolate between this regime and the normal gravitational instability criterion using $T_{\epsilon} = (T_{h \gg \epsilon}^{-1} + T_{\epsilon \gg h}^{-1})^{-1}$. This ensures a smooth transition between the two limiting cases. The left panel of Fig. 18 shows that the density

above which the softened instability criterion starts to deviate from the unsoftened criterion corresponds to $n_{\text{H,soft}}$ as expected.

Gas below the dashed curves in the left panel of Fig. 18 is gravitationally unstable⁹, while gas between the solid and dashed curves should be unstable, but is stabilized by gravitational softening. The thermal equilibrium temperatures (dotted lines) are located mostly below the dashed curves, which suggests that in practice gravitational softening will generally not directly prevent SF (assuming $\sigma_{\text{th}} \gg \sigma_{\text{turb}}$). For our fiducial resolution, $M_{\text{particle}} = 10^5 M_{\odot}$, the equilibrium temperature of metal-free gas does extend in the problematic regime between the solid and dashed curves and the same is true for low-metallicity gas at very high densities, $n_{\text{H}} > 10^4 \text{ cm}^{-3}$. However, these problematic densities exceed $n_{\text{H,soft}}$ by many orders of magnitude, which means self-gravity is unlikely to produce them, though they may occur in pressure-confined clouds and in the centres of clouds with mass $\gg \langle N_{\text{ngb}} \rangle M_{\text{particle}}$ owing to the weight of the overlying layers.

For fixed $M_{\text{particle}}/\epsilon^3$, lower-resolution simulations, i.e. simulations using a larger M_{particle} , are able to resolve gravitational instabilities up to higher densities (i.e., the dashed instability curves in the left panel of Fig. 18 shift to higher densities for a fixed temperature). This may sound counterintuitive, but it is a consequence of the fact that we evaluate the instability criterion at the resolution limit, i.e. for a cloud mass that scales with the particle mass. This suggests that keeping $M_{\text{particle}}/\epsilon^3$ fixed may not be desirable when the particle mass is decreased to small values.

In the right panel of Fig. 18 we again show the unsoftened and softened gravitational instability criteria in temperature-density space for our fiducial resolution and assuming $\sigma_{\text{th}} \gg \sigma_{\text{turb}}$ and compare it with gravitational instability criteria for smaller gravitational softening lengths, keeping the particle mass fixed at our fiducial value. This shows that when the gravitational softening length is a factor of two smaller, the softened gravitational instability curve begins to deviate from the line corresponding to the unsoftened gravitational softening length at a factor eight higher densities. This is the opposite trend as was found by Ploekinger et al. in prep. for artificial collapse triggered by the minimum SPH smoothing length. However, the impact at high densities is larger because equation (48) scales strongly with ϵ and we obtain a softening dependent density at which 10 K gas is unstable given by

$$n_{\text{H},\epsilon} < 3.0 \times 10^4 \text{ cm}^{-3} \left(\frac{T}{10 \text{ K}} \right)^{-3/2} \left(\frac{\epsilon}{200 \text{ pc}} \right)^{-9/2}. \quad (54)$$

This means that a factor of two in the gravitational softening length at fixed particle mass will change the maximal density for which the gas is unstable by 1.35 dex.

Fig. 19 shows the t_{gas} relation for simulations with different gravitational softening lengths. The t_{gas} relations depend only weakly on the gravitational softening lengths. The break in the relations shifts to smaller molecular surface densities for the largest gravitational softening length (800 pc). At high surface densities the gas consumption time scales increase systematically with the softening length, but the effect is small compared with the scatter.

Altogether we find that a softened or unsoftened gravitational instability criterion selects almost the same gas when the softening length is small enough for the mass resolution. The gas considered stable in softened gravity is mainly at high densities and higher temperatures than the equilibrium temperatures. Because of this,

⁹ However, it has a longer free-fall time than the Newtonian free-fall time, so the collapse takes longer and follows equation (38).

the impact on the star-forming gas is limited and the KS relations are not influenced much.

If we want to resolve the same region in the $\rho - T$ plane, we can use equation (48) to give a relation between the particle resolution and the gravitational softening length of

$$\epsilon = 200 \text{ pc} \left(\frac{M_{\text{particle}}}{10^5 M_{\odot}} \right)^{5/9}, \quad (55)$$

which has a stronger dependence on the particle mass as compared to the typical scaling (i.e. $\epsilon \propto M_{\text{particle}}^{1/3}$). However, this still does not allow us to resolve higher density than the limit given by equation (54). If we use a scaling of the gravitational softening length given by

$$\epsilon = 200 \text{ pc} \left(\frac{M_{\text{particle}}}{10^5 M_{\odot}} \right)^n, \quad (56)$$

then it is required that $n > 5/9$ to resolve higher densities correctly for higher resolutions. A slight increase gives a value of $n = 2/3$. This would always resolve the same region of $\rho - T$ and would allow us to resolve higher densities when we have higher resolution. However, using similar small gravitational softening lengths for stellar particles could result in undesired artificial scattering of stellar particles (Ludlow et al. 2019, 2020, 2021; Wilkinson et al. 2022).

5.2 Comparison with the literature

This is not the first attempt to understand the SF scaling relations in hydrodynamical simulations. However, many impose an equation of state to prevent the gas from cooling down to low ($T \ll 10^4 \text{ K}$) temperatures (e.g. Springel & Hernquist 2003; Stinson et al. 2006; Schaye & Dalla Vecchia 2008). Marinacci et al. (2019) showed that there are significant differences between hydrodynamical simulations with and without an equation of state, such as a stronger variation of the SFH on time scales of $t \sim 20 \text{ Myr}$ and a larger scatter in Σ_{SFR} at fixed gas surface densities. We therefore focus on comparing our results with simulations that explicitly model cold gas ($T \lesssim 10^2 \text{ K}$).

Simulations of idealised disk galaxies indicate that the SFH is insensitive to the SF criterion. This is the case for a wide variation of SF criteria, such as density thresholds, temperature ceilings, converging flows, $t_{\text{cool}} < t_{\text{ff}}$, gravitational instability and combinations thereof (e.g. Hopkins et al. 2013; Sillero et al. 2021; and our Fig. 14) and it is also true for cosmological simulations (e.g. Hopkins et al. 2018, 2023) and cosmological simulations that use an equation of state (Schaye et al. 2010). The SF criterion does however have a significant effect on the physical conditions of the gas from which stars form (e.g. Hopkins et al. 2013, 2023) and very strict criteria, like density thresholds that are high relative to the numerical resolution, can produce artificial features (e.g. Fig. 6 of Hopkins et al. 2023). Furthermore, the combination of multiple SF criteria is dominated by the most strict SF criterion (e.g. a density threshold in FIRE2 and a gravitational instability criterion in FIRE3, see Figs. 6 and A1 of Hopkins et al. 2023). Becerra & Escala (2014) found that with higher density thresholds the normalisation of the total t_{gas} relation decreases. This is opposite to what we are finding (see Fig. 15). While a higher density threshold leads to stars forming at higher surface densities and hence from gas with shorter consumption time scales, the fraction of gas that is star forming decreases, which leads to longer gas consumption time scales. Apparently, in our simulations the latter effect wins, yielding higher normalisations of the t_{gas} relations for higher density thresholds.

Matching the observed high-surface density slope of the spatially-resolved and azimuthally-averaged neutral gas KS relation ($N_{\text{KS}} \approx 1.4$) or, equivalently, the slope of the t_{gas} relation ($N \approx -0.4$) is challenging. Some simulations agree with the canonical values (e.g. [Becerra & Escala 2014](#); [Semenov et al. 2016, 2017](#); [Sillero et al. 2021](#)), while we ($N \approx -0.7$, see Fig. 4) and others predict relations that are slightly steeper ($N \approx -0.7$, [Orr et al. 2018](#)). Other simulations predict relations that are so steep that they do not reproduce the observations ($N \approx -1.5$, [Agertz & Kravtsov 2015](#); $N \approx -3$ to -5 , [Marinacci et al. 2019](#); $N \approx -1.5$, [Gensior et al. 2020](#); $N \approx -1$ to -4 , [Bieri et al. 2023](#)).

The main ingredients responsible for the normalisation of the t_{gas} relation are the SFE and the radiative cooling. Prescriptions for radiative cooling that do not form enough molecular gas decrease the normalisation of the molecular t_{gas} relation (similar to the metallicity trend of Fig. 4). This seems to be the case for [Semenov et al. \(2017\)](#), who found that their molecular t_{gas} relation has a 0.5 dex lower normalisation than observed, their neutral t_{gas} relation agrees with the data. The SFE does not change the average SFR significantly for values between 1 and 100 per cent (e.g. [Li et al. 2020](#)), but it can significantly change the normalisation of the t_{gas} relation as shown in Fig. 16. The FIRE2 simulations ([Hopkins et al. 2018](#)) use an SFE of 100 per cent, which results in a molecular t_{gas} relation with a 1.5 dex too low normalisation (Fig. 2 of [Orr et al. 2018](#)). [Agertz et al. \(2013\)](#) showed that a 10 per cent SFE gives a too low normalisation while 1 per cent is in agreement with observations. Similarly, [Semenov et al. \(2018\)](#) showed that a higher SFE decreases the normalisation significantly (their Fig. 10) and an SFE of 1 per cent is in agreement with the data. These results are all in agreement with what we find. This suggests that using a low SFE of ~ 1 per cent is required because we are still far away from resolving the densities at which actual stars form and the efficiency is close to 100 per cent.

5.3 Which SF criterion is best?

In §4.2.1 we showed that the spatially-resolved t_{gas} relation is an excellent discriminator of SF criteria. Based on Fig. 15 alone, we find that at the resolution studied in this paper and for solar metallicity, the gravitational instability criterion (for both $\alpha_{\text{crit}} = 1$ and 3) and the temperature criterion (for $T_{\text{crit}} \leq 10^3$ K) all reproduce the observed t_{gas} relations (although the temperature criterion requires an adjusted SFE). However, as shown in Fig. 9, a fixed criterion in ρ - T space, which for gravitational instability at the resolution limit implies a value of α_{crit} that depends on the numerical resolution, results in poor convergence of the t_{gas} relation with the numerical resolution. Instead, a gravitational instability criterion with a fixed value of α_{crit} , which corresponds to different regions in ρ - T space for different numerical resolutions, reproduces the t_{gas} relation and converges much better with the numerical resolution.

For low-metallicity gas, the behaviour is significantly different for a temperature ceiling and a gravitational instability criterion. Very low metallicity gas requires very high densities to cool below temperatures of $T \approx 10^3$ K, particularly if an interstellar radiation field is included. This means that cosmological simulations with a temperature criterion that do not resolve such high densities might be unable to form any stars at all. A gravitational instability criterion with $\alpha_{\text{crit}} \sim 1$ will however enable SF at higher temperatures in regions with sufficiently low turbulent velocity dispersion and sufficiently high densities.

While stars form in cold molecular clouds, generally not all cold gas is gravitationally unstable. Cold gas with a large turbulent

dispersion will not be gravitationally unstable (see Fig. 1) and hence should not be considered star forming. Therefore, situations that can cause an increase in the turbulent velocity, like the presence of a bulge or stellar feedback, will make it harder to form stars, e.g. leading to morphological quenching (e.g. [Martig et al. 2009, 2013](#)) or self-regulation. A temperature criterion that just determines if gas is cold is less sensitive to such effects (though supersonic turbulence may be converted to heat).

Overall, these points indicate that a gravitational instability criterion (with a fixed $\alpha_{\text{crit}} \sim 1$) is the better choice for a SF criterion.

5.4 Caveats

The gas and stellar disks in our idealised disk galaxies are initialized with a constant density at a fixed radius. This means that any spiral structure in the gas is developed by internally triggered instabilities. However, spiral structure can develop due to interactions with satellite galaxies (see [Dobbs & Baba 2014](#), for a review), a nearby galaxy cluster (e.g. [Semczuk et al. 2017](#)), asymmetries in the DM distribution ([Khoperskov et al. 2013](#)), or tidal perturbations from DM subhaloes (e.g. [Chang & Chakrabarti 2011](#)). [Pettitt et al. \(2020\)](#) showed that including a satellite galaxy does impact the structure of the ISM at large radii ($r > 3$ kpc) producing higher densities in the arms and lower densities between them. Similarly, the initial gas fraction and metallicity are constant throughout the disk, which is unlikely to be true for real galaxies.

Furthermore, our simulations lack a circumgalactic medium (CGM), the presence of cosmological accretion and a cosmological environment. This means outflows are able to escape with higher velocities and to larger distances than would be the case if a CGM were included and that there is no replenishment of gas consumed by SF. Therefore a limitation of idealised galaxy simulations is that they are probably unsuitable for investigating the spatially-resolved SF main sequence and gas main sequence.

Our adopted gas fraction of 30 per cent is high for galaxies observed at $z = 0$. However, in order to test SF models, we need a high enough $\Sigma_{\text{H}+\text{H}_2}$ to reach the power-law regime of the KS relations. A lower gas fraction would increase the stellar surface density at fixed $\Sigma_{\text{H}+\text{H}_2}$, which may push the break in the neutral KS relation to slightly lower $\Sigma_{\text{H}+\text{H}_2}$. Note that this would improve the agreement with the data.

Our disk galaxy does not contain a classical bulge component. Real galaxies and galaxies formed in cosmological simulations contain bulge components that vary in strength and are formed by the collisions of galaxies or disk instabilities (see [Brooks & Christensen 2016](#), for a review). The inclusion of a bulge component could suppress SF in the centre (e.g. [Martig et al. 2009, 2013](#)). We leave an investigation of the effect of the bulge for future work.

Lastly, our simulations use a fixed gravitational softening length ϵ , while an adaptive gravitational softening length is often used (e.g. [Price & Monaghan 2007](#) and models based on this like [Springel 2010](#); [Iannuzzi & Dolag 2011](#)). However, as we show in Fig. 19 different gravitational softening lengths give almost identical results.

6 CONCLUSIONS

We have used SPH simulations of isolated Milky Way mass disk galaxies that include cold, interstellar gas to test subgrid prescriptions for star formation (SF). Our fiducial model consists of a gravi-

tational instability criterion that includes both thermal and turbulent motions and is evaluated at the mass resolution limit (equation 27 with $\alpha_{\text{crit}} = 1$) combined with the Schmidt SF law (equation 1) with an efficiency per free fall time of $\varepsilon = 0.01$. Our fiducial simulation assumes solar metallicity and uses a particle mass of $10^5 M_{\odot}$ and a gravitational softening length of 200 pc. We investigated the impact of the numerical resolution, SF criterion, subgrid SF efficiency, supernova feedback energy, metallicity, and the binning of the observables on the predicted Kennicutt-Schmidt (KS) SF law for neutral, molecular and atomic gas. The main conclusions we draw are:

(i) Our fiducial model matches the observed neutral, molecular and atomic Kennicutt-Schmidt (KS) relations well for both spatially-resolved (Fig. 4) and azimuthally-averaged data (Fig. 6).

(ii) Varying the spatial bin size between 100 pc and 10 kpc shows that the spatial averaging scale strongly affects the spatially-resolved KS relations. As expected, the scatter decreases with the spatial bin size. At low surface densities a smaller bin size results in a longer median gas consumption time scale. However, for a given bin size there is a surface density above which the neutral and molecular, but not the atomic, KS laws converge to the corresponding KS laws for larger bin sizes (Fig. 5). Azimuthally-averaged KS laws using a bin size similar to the bin size used in observations (~ 750 pc) resemble spatially-resolved KS laws using larger bin sizes (similar to the disk scale length) (Fig. 6).

(iii) Decreasing the metallicity shifts the break in the neutral gas KS relation to somewhat higher surface densities. For the molecular KS law the effect is more dramatic, with lower metallicities yielding shorter gas consumption time scales at fixed molecular surface density because the molecular density more strongly underestimates the true gas density (Fig. 13).

(iv) Observational and theoretical studies of SF laws generally ignore ionized gas. However, for neutral hydrogen surface densities $\ll 1 M_{\odot} \text{pc}^{-2}$ ionized gas dominates the total gas surface density and at a neutral surface density of $10 M_{\odot} \text{pc}^{-2}$ ionized gas still contributes about as much as molecular hydrogen does (for a spatial averaging scale of 750 pc; Fig. 10).

(v) To investigate the numerical convergence, we vary the mass (spatial) resolution by a factor of 4096 (16). We find good convergence for the SF history (Fig. 7), for the neutral KS law, and at $\Sigma_{\text{H}_2} \gg 10 M_{\odot} \text{pc}^{-2}$ also for the molecular KS law (Fig. 9). Even though we show that gravitational softening does modify the region of $\rho - T$ space that is unstable (Fig. 18), we find that the KS laws are largely insensitive to the gravitational softening length (when varied between 50 and 800 pc at our fiducial mass resolution). Although at very high surface densities ($\gtrsim 10^2 M_{\odot} \text{pc}^{-2}$) the gas consumption time scale increases slightly with the softening length (Fig. 19).

(vi) The gravitational instability criterion selects different regions of $T - n_{\text{H}}$ plane depending on the mass resolution and the value of α_{crit} . Higher (lower) mass resolution means that gas with higher (lower) densities and lower (higher) temperatures is selected to be star-forming (Fig. 1). While this resolution dependence can largely be circumvented by scaling α_{crit} with the resolution (Fig. 8), we think it is desirable to evaluate the SF criterion at the simulation's resolution limit, because it allows higher-resolution models to directly simulate the physics up to higher densities. We thus argue that numerical convergence is desirable for observables, but not for the physical properties of gas that is labelled star-forming by the subgrid model. In fact, we find that a fixed value of α_{crit} gives better convergence for the KS laws than a value of α_{crit} that is scaled with the resolution (Fig. 9).

(vii) We compare simulations using the gravitational instability criterion for different values of α_{crit} with simulations using a range of density thresholds or temperature ceilings. While SF histories are insensitive to the SF criterion (Fig. 14), the KS relations can discriminate between different models (Fig. 15). Comparison with the data rules out gravitational instability criteria with $\alpha_{\text{crit}} \ll 1$ or $\alpha_{\text{crit}} \gg 1$ and also all the models with a temperature ceiling or density threshold. Additionally, we prefer the instability criterion over a temperature ceiling because it is more robust for cosmological simulations, where (nearly) metal-free gas may only cool to low temperatures at very high densities that typically remain unresolved, and because physically we do not expect all cold gas to form stars.

(viii) Varying the SF efficiency per free-fall time by a factor of 10^3 shows that it has little effect on the SF history, but has a strong impact on the normalization of and the break in the KS relations. Comparison with observations constrains the SF efficiency to be of order 1 per cent (Fig. 16).

(ix) Varying the strength of stellar feedback by a factor of 32 shows that it has a large impact on the SF history and a non-negligible effect on the normalization of the KS laws, particularly for the neutral gas (Fig. 17).

In future work we plan to use our fiducial SF prescription in cosmological simulations of galaxy formation. This will enable us to include effects such as gas accretion and galaxy interactions, and to investigate a wide range of galaxy masses, gas fractions, and morphologies.

ACKNOWLEDGEMENTS

We would like to thank Sara Ellison, Ismael Pessa, Miguel Querejeta and Andreas Schruha for kindly providing us their KS relation data points for comparing with our simulations. The research in this paper made use of the SWIFT open-source simulation code (<http://www.swiftsim.com>, Schaller et al. 2018) version 0.9.0. This work used swiftsimio (Borrow & Borisov 2020) for reading and visualising the data. SPH projections use the subsampled projection backends (Borrow & Kelly 2021). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. EC is supported by the funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860744 (BiD4BEST).

DATA AVAILABILITY

The initial conditions of the simulations presented in this paper are publicly available in SWIFT. SWIFT is publicly available and can be found at www.swiftsim.com. The other data underlying this article will be shared on a reasonable request to the corresponding author.

REFERENCES

- Abdurro'uf Lin Y.-T., Hirashita H., Morishita T., Tacchella S., Wu P.-F., Akiyama M., Takeuchi T. T., 2022, *ApJ*, **935**, 98
 Agertz O., Kravtsov A. V., 2015, *ApJ*, **804**, 18

- Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, *ApJ*, **770**, 25
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, **47**, 481
- Becerra F., Escala A., 2014, *ApJ*, **786**, 56
- Bieri R., Naab T., Geen S., Coles J. P., Pakmor R., Walch S., 2023, *MNRAS*, **523**, 6336
- Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, **136**, 2846
- Bigiel F., Leroy A., Walter F., Blitz L., Brinks E., de Blok W. J. G., Madore B., 2010, *AJ*, **140**, 1194
- Bolato A. D., et al., 2017, *ApJ*, **846**, 159
- Borrow J., Borrisov A., 2020, *Journal of Open Source Software*, **5**, 2430
- Borrow J., Kelly A. J., 2021, arXiv e-prints, p. arXiv:2106.05281
- Borrow J., Schaller M., Bower R. G., Schaye J., 2022, *MNRAS*, **511**, 2367
- Borrow J., Schaller M., Bahé Y. M., Schaye J., Ludlow A. D., Ploeckinger S., Nobels F. S. J., Altamura E., 2023, *MNRAS*, **526**, 2441
- Bosma A., 1981, *AJ*, **86**, 1825
- Broeils A. H., van Woerden H., 1994, *A&AS*, **107**, 129
- den Brok J. S., et al., 2021, *MNRAS*, **504**, 3221
- den Brok J. S., et al., 2023, arXiv e-prints, p. arXiv:2302.03044
- Brooks A., Christensen C., 2016, in Laurikainen E., Peletier R., Gadotti D., eds, *Astrophysics and Space Science Library* Vol. 418, Galactic Bulges. p. 317 (arXiv:1511.04095), doi:10.1007/978-3-319-19378-6_12
- Bundy K., et al., 2015, *ApJ*, **798**, 7
- Chabrier G., 2003, *PASP*, **115**, 763
- Chaikin E., Schaye J., Schaller M., Bahé Y. M., Nobels F. S. J., Ploeckinger S., 2022, *MNRAS*, **514**, 249
- Chaikin E., Schaye J., Schaller M., Benítez-Llambay A., Nobels F. S. J., Ploeckinger S., 2023, *MNRAS*, **523**, 3709
- Chang P., Chakrabarti S., 2011, *MNRAS*, **416**, 618
- Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, *MNRAS*, **452**, 1217
- Courant R., Friedrichs K., Lewy H., 1928, *Mathematische Annalen*, **100**, 32
- Cullen L., Dehnen W., 2010, *MNRAS*, **408**, 669
- Dalla Vecchia C., Schaye J., 2012, *MNRAS*, **426**, 140
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *MNRAS*, **486**, 2827
- Dessauges-Zavadsky M., Verdugo C., Combes F., Pfneniger D., 2014, *A&A*, **566**, A147
- Dobbs C., Baba J., 2014, *Publ. Astron. Soc. Australia*, **31**, e035
- Dubois Y., et al., 2021, *A&A*, **651**, A109
- Durier F., Dalla Vecchia C., 2012, *MNRAS*, **419**, 465
- Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, *Publ. Astron. Soc. Australia*, **34**, e058
- Ellison S. L., et al., 2020, *MNRAS*, **493**, L39
- Evans II N. J., Heiderman A., Vutisalchavakul N., 2014, *ApJ*, **782**, 114
- Faucher-Giguère C.-A., 2020, *MNRAS*, **493**, 1614
- Feldmann R., et al., 2023, *MNRAS*, **522**, 3831
- Fogarty L. M. R., et al., 2015, *MNRAS*, **454**, 2050
- Genel S., et al., 2019, *ApJ*, **871**, 21
- Gensior J., Kruijssen J. M. D., Keller B. W., 2020, *MNRAS*, **495**, 199
- Greengard L., Rokhlin V., 1987, *Journal of Computational Physics*, **73**, 325
- Grisdale K., 2021, *MNRAS*, **500**, 3552
- Grisdale K., Agertz O., Renaud F., Romeo A. B., Devriendt J., Slyz A., 2019, *MNRAS*, **486**, 5482
- Grudić M. Y., Hopkins P. F., Lee E. J., Murray N., Faucher-Giguère C.-A., Johnson L. C., 2019, *MNRAS*, **488**, 1501
- Heiderman A., Evans Neal J. I., Allen L. E., Huard T., Heyer M., 2010, *ApJ*, **723**, 1019
- Hernquist L., 1990, *ApJ*, **356**, 359
- Hopkins P. F., Narayanan D., Murray N., 2013, *MNRAS*, **432**, 2647
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, **445**, 581
- Hopkins P. F., et al., 2018, *MNRAS*, **480**, 800
- Hopkins P. F., et al., 2023, *MNRAS*, **519**, 3154
- Hu C.-Y., Naab T., Walch S., Glover S. C. O., Clark P. C., 2016, *MNRAS*, **458**, 3528
- Hu Z., Krumholz M. R., Pokhrel R., Gutermuth R. A., 2022, *MNRAS*, **511**, 1431
- Iannuzzi F., Dolag K., 2011, *MNRAS*, **417**, 2846
- Indriolo N., et al., 2015, *ApJ*, **800**, 40
- Jeans J. H., 1902, *Philosophical Transactions of the Royal Society of London Series A*, **199**, 1
- Keller B. W., Wadsley J. W., Wang L., Kruijssen J. M. D., 2019, *MNRAS*, **482**, 2244
- Kennicutt Jr. R. C., 1989, *ApJ*, **344**, 685
- Kennicutt Jr. R. C., 1998, *ApJ*, **498**, 541
- Kennicutt Jr. R. C., et al., 2007, *ApJ*, **671**, 333
- Khoperskov A. V., Khoperskov S. A., Zasov A. V., Bizyaev D. V., Khrapov S. S., 2013, *MNRAS*, **431**, 1230
- Koopmann R. A., Kenney J. D. P., 2004, *ApJ*, **613**, 851
- Kretschmer M., Teyssier R., 2020, *MNRAS*, **492**, 1385
- Larson R. B., 1981, *MNRAS*, **194**, 809
- Lee E. J., Miville-Deschênes M.-A., Murray N. W., 2016, *ApJ*, **833**, 229
- Leroy A. K., et al., 2017, *ApJ*, **846**, 71
- Leroy A. K., et al., 2021, *ApJS*, **255**, 19
- Leroy A. K., et al., 2022, *ApJ*, **927**, 149
- Li H., Vogelsberger M., Marinacci F., Sales L. V., Torrey P., 2020, *MNRAS*, **499**, 5862
- Lin L., et al., 2019, *ApJ*, **884**, L33
- Ludlow A. D., Schaye J., Schaller M., Richings J., 2019, *MNRAS*, **488**, L123
- Ludlow A. D., Schaye J., Schaller M., Bower R., 2020, *MNRAS*, **493**, 2926
- Ludlow A. D., Fall S. M., Schaye J., Obreschkow D., 2021, *MNRAS*, **508**, 5114
- Maoz D., Mannucci F., 2012, *Publ. Astron. Soc. Australia*, **29**, 447
- Maoz D., Sharon K., Gal-Yam A., 2010, *ApJ*, **722**, 1879
- Marinacci F., Sales L. V., Vogelsberger M., Torrey P., Springel V., 2019, *MNRAS*, **489**, 4233
- Martig M., Bournaud F., Teyssier R., Dekel A., 2009, *ApJ*, **707**, 250
- Martig M., et al., 2013, *MNRAS*, **432**, 1914
- Martin C. L., Kennicutt Robert C. J., 2001, *ApJ*, **555**, 301
- Mo H. J., Mao S., White S. D. M., 1998, *MNRAS*, **295**, 319
- Murphy E. J., et al., 2015, *ApJ*, **813**, 118
- Murray N., 2011, *ApJ*, **729**, 133
- Narayanan D., Krumholz M. R., Ostriker E. C., Hernquist L., 2012, *MNRAS*, **421**, 3127
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, **490**, 493
- Nelson D., et al., 2019, *MNRAS*, **490**, 3234
- Nobels F. S. J., Schaye J., Schaller M., Bahé Y. M., Chaikin E., 2022, *MNRAS*, **515**, 4838
- Ochsendorf B. B., Meixner M., Roman-Duval J., Rahman M., Evans Neal J. I., 2017, *ApJ*, **841**, 109
- Onodera S., et al., 2010, *ApJ*, **722**, L127
- Oppenheimer B. D., 2018, *MNRAS*, **480**, 2963
- Orr M. E., et al., 2018, *MNRAS*, **478**, 3653
- Padoan P., Haugbølle T., Nordlund Å., 2012, *ApJ*, **759**, L27
- Pessa I., et al., 2021, *A&A*, **650**, A134
- Pettitt A. R., Dobbs C. L., Baba J., Colombo D., Duarte-Cabral A., Egusa F., Habe A., 2020, *MNRAS*, **498**, 1159
- Pillepich A., et al., 2018, *MNRAS*, **473**, 4077
- Planck Collaboration et al., 2016, *A&A*, **594**, A13
- Ploeckinger S., Schaye J., 2020, *MNRAS*, **497**, 4857
- Ploeckinger S., Nobels F. S. J., Schaller M., Schaye J., 2024, *MNRAS*, **528**, 2930
- Plummer H. C., 1911, *MNRAS*, **71**, 460
- Pokhrel R., et al., 2021, *ApJ*, **912**, L19
- Portinari L., Chiosi C., Bressan A., 1998, *A&A*, **334**, 505
- Price D. J., 2012, *Journal of Computational Physics*, **231**, 759
- Price D. J., Monaghan J. J., 2007, *MNRAS*, **374**, 1347
- Querejeta M., et al., 2019, *A&A*, **625**, A19
- Querejeta M., et al., 2021, *A&A*, **656**, A133
- Raskutti S., Ostriker E. C., Skinner M. A., 2016, *ApJ*, **829**, 130
- Reeves S. N., Sadler E. M., Allison J. R., Koribalski B. S., Curran S. J., Pracy M. B., 2015, *MNRAS*, **450**, 926

- Revaz Y., Jablonka P., 2018, *A&A*, 616, A96
 Richings A. J., Schaye J., 2016, *MNRAS*, 458, 270
 Richings A. J., Schaye J., Oppenheimer B. D., 2014a, *MNRAS*, 440, 3349
 Richings A. J., Schaye J., Oppenheimer B. D., 2014b, *MNRAS*, 442, 2780
 Sánchez S. F., et al., 2012, *A&A*, 538, A8
 Schaller M., Gonnet P., Chalk A. B. G., Draper P. W., 2016, in Proceedings of the PASC Conference. PASC '16. ACM, New York, NY, USA (arXiv:1606.02738), doi:10.1145/2929908.2929916
 Schaller M., Gonnet P., Draper P. W., Chalk A. B. G., Bower R. G., Willis J., Hausammann L., 2018, SWIFT: SPH With Inter-dependent Fine-grained Tasking, Astrophysics Source Code Library (ascl:1805.020)
 Schaller M., et al., 2023, arXiv e-prints, p. arXiv:2305.13380
 Schaye J., 2004, *ApJ*, 609, 667
 Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210
 Schaye J., et al., 2010, *MNRAS*, 402, 1536
 Schaye J., et al., 2015, *MNRAS*, 446, 521
 Schmidt M., 1959, *ApJ*, 129, 243
 Schöier F. L., van der Tak F. F. S., van Dishoeck E. F., Black J. H., 2005, *A&A*, 432, 369
 Schrubba A., et al., 2011, *AJ*, 142, 37
 Semczuk M., Łokas E. L., del Pino A., 2017, *ApJ*, 834, 7
 Semenov V. A., Kravtsov A. V., Gnedin N. Y., 2016, *ApJ*, 826, 200
 Semenov V. A., Kravtsov A. V., Gnedin N. Y., 2017, *ApJ*, 845, 133
 Semenov V. A., Kravtsov A. V., Gnedin N. Y., 2018, *ApJ*, 861, 4
 Semenov V. A., Kravtsov A. V., Gnedin N. Y., 2019, *ApJ*, 870, 79
 Sillero E., Tissera P. B., Lambas D. G., Bovino S., Schleicher D. R., Grassi T., Bruzual G., Charlot S., 2021, *MNRAS*, 504, 2325
 Springel V., 2010, *MNRAS*, 401, 791
 Springel V., Hernquist L., 2003, *MNRAS*, 339, 289
 Springel V., Di Matteo T., Hernquist L., 2005, *MNRAS*, 361, 776
 Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
 Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, *MNRAS*, 373, 1074
 Strömgren B., 1939, *ApJ*, 89, 526
 Teng Y.-H., et al., 2022, *ApJ*, 925, 72
 Utomo D., et al., 2018, *ApJ*, 861, L18
 Viaene S., Forbrich J., Fritz J., 2018, *MNRAS*, 475, 5550
 Vutisalchavakul N., Evans II N. J., Heyer M., 2016, *ApJ*, 831, 73
 Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt Robert C. J., Thornley M. D., Leroy A., 2008, *AJ*, 136, 2563
 Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, *MNRAS*, 454, 83
 Whitworth D. J., Smith R. J., Tress R., Kay S. T., Glover S. C. O., Sormani M. C., Klessen R. S., 2022, *MNRAS*, 510, 4146
 Wilkinson M. J., Ludlow A. D., Lagos C. d. P., Fall S. M., Schaye J., Obreschkow D., 2022, arXiv e-prints, p. arXiv:2208.07623
 Yajima Y., et al., 2021, *PASJ*, 73, 257

APPENDIX A: METALLICITY VARIATIONS TEMPERATURE CRITERION

Fig. A1 shows the metallicity dependence of the KS relation for a temperature criterion ($T < 10^3$ K). The dependence on metallicity is slightly stronger than for a gravitational instability criterion.

APPENDIX B: STAR FORMATION HISTORIES FOR DIFFERENT VARIATIONS

Fig. B1 shows the 100 Myr moving average of the SFR for different SFE. In general the SFH is affected mainly in the start by the SFE. At later times, a $\epsilon = 1.0$ has a lower value because it produces more outflows that deplete the galaxy of gas.

Fig. B2 shows the 100 Myr moving average of the SFR for

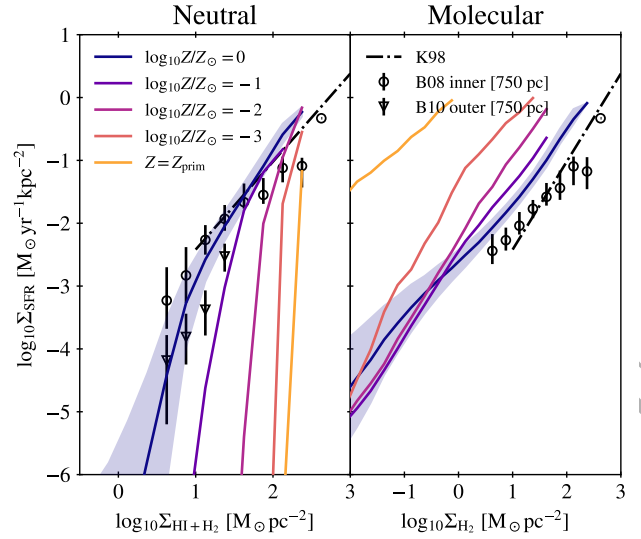


Figure A1. Comparison of the resolved (750 pc) neutral and molecular KS relation at different metallicities (different colours) for the temperature criterion. The shaded region shows the 16th and 84th percentile scatter for the solar metallicity simulation. We show the disk averaged KS relation from K98. The neutral KS relation changes almost two dex in the position of the break while the molecular KS relation has a change in normalization of more than two orders of magnitude compared to solar metallicity.

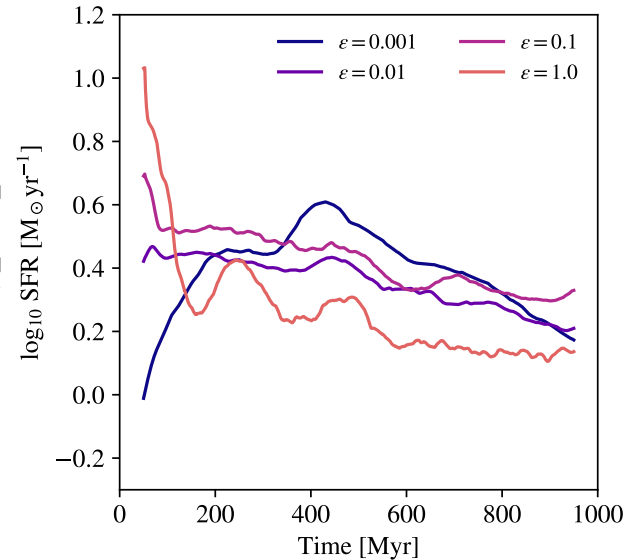


Figure B1. Comparison of the star formation histories for the simulations with different SFE (different colours). The lines show the 100 Myr moving average of the SFR. The differences are small and are mainly.

different CC SN energies. A larger CC SNe energy results in a lower SFR.

Fig. B3 shows the 100 Myr moving average of the SFR for different metallicities comparing primordial to solar metallicity. Lower metallicities correspond to lower SFR, for and below metallicities of 1 per cent the SFH is mostly unchanged.

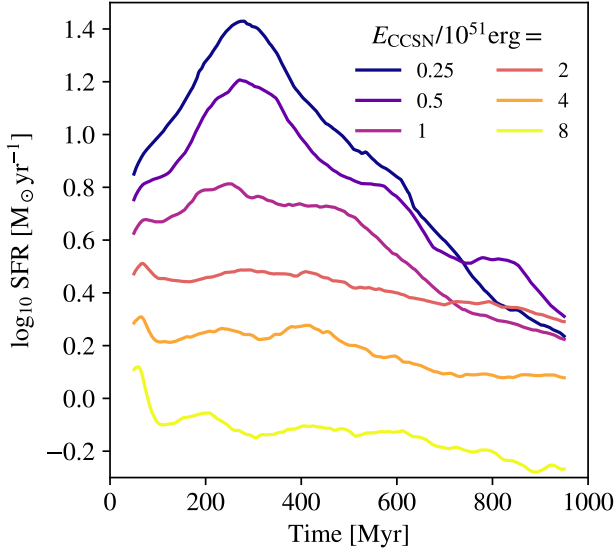


Figure B2. Comparison of the star formation histories for the simulations with different CC SN energies (different colours). The lines show the 100 Myr moving average of the SFR.

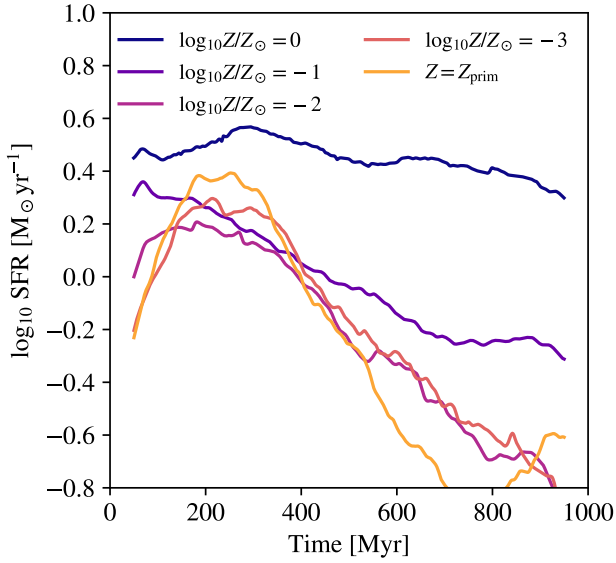


Figure B3. Comparison of the star formation histories for the simulations with different metallicities (different colours). The lines show the 100 Myr moving average of the SFR.