

Clouds of Theseus: long-lived molecular clouds are composed of short-lived H₂ molecules

Sarah M. R. Jeffreson ¹★, Vadim A. Semenov ¹ and Mark R. Krumholz ^{2,3}

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

²Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

³Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Canberra, ACT 2611, Australia

Accepted 2023 November 13. Received 2023 November 13; in original form 2023 January 26

ABSTRACT

We use passive gas tracer particles in an AREPO simulation of a dwarf spiral galaxy to relate the Lagrangian evolution of star-forming gas parcels and their H₂ molecules to the evolution of their host giant molecular clouds. We find that the median chemical lifetime of H₂ is 4 Myr, with an interquartile range between 2 and 9 Myr. This chemical lifetime is independent of the lifetime of the host molecular cloud, which may extend up to 90 Myr, with around 50 per cent of star formation occurring in longer lived clouds (>25 Myr). The rapid ejection of gas from around young massive stars by early stellar feedback is responsible for the short H₂ survival time, driving down the density of the surrounding gas, so that its H₂ molecules are dissociated by the interstellar radiation field. This ejection of gas from the H₂-dominated state is balanced by the constant accretion of new gas from the galactic environment, constituting a ‘competition model’ for molecular cloud evolution. Gas ejection occurs at a rate that is proportional to the molecular cloud mass, so that the cloud lifetime is determined by the accretion rate, which may be as high as $4 \times 10^4 M_{\odot} \text{ Myr}^{-1}$ in the longest lived clouds. Our findings therefore resolve the conflict between observations of rapid gas ejection around young massive stars and observations of long-lived molecular clouds in galaxies. We show that the fastest-accreting, longest lived, highest mass clouds drive supernova clustering on sub-cloud scales, which in turn is a key driver of galactic outflows.

Key words: ISM: clouds – ISM: evolution – ISM: structure – galaxies: star formation.

1 INTRODUCTION

In recent years, a new, dynamical picture of star formation has emerged. Numerical simulations of disc galaxies have demonstrated that Lagrangian parcels of gas in the interstellar medium undergo constant and rapid cycles of compression into, and disruption out of, a high-density ‘star-forming state’ (Semenov, Kravtsov & Gnedin 2017, 2018; Shin et al. 2022). The average time spent by gas in the star-forming state (~ 5 –15 Myr) is much shorter than the time spent outside the star-forming state ($\gtrsim 100$ Myr). This picture can explain several of the key observed attributes of star formation in disc galaxies. The fact that galaxies convert only a small fraction of their gas to stars per galactic rotation (e.g. Kennicutt 1998; Wyder et al. 2009; Daddi et al. 2010) is consistent with the small fraction of time spent by gas in the star-forming state. The observed spread of star formation rates (SFRs) at a given molecular gas surface density found when galaxies are observed at high resolution (Onodera et al. 2010; Schrubba et al. 2011), and the relatively poor correlation between molecular masses and SFR indicators for clouds in the Milky Way (Mooney & Solomon 1988; Lee, Miville-Deschênes & Murray 2016), reflect the fact that any particular parcel of H₂-rich gas may have just started forming stars (in which case its luminosity per unit

gas mass in SFR tracers such as H α or infrared emission will be low) or may have just completed a star formation cycle and left the star-forming state (in which case the luminosity per unit mass will be high). The spatial decorrelation of molecular gas and young stars on small scales ($\lesssim 100$ pc) relative to galactic scales (~ 1 kpc, Schrubba et al. 2010; Kruijssen et al. 2019) and the rapidity with which star clusters become optically revealed (Hollyhead et al. 2015; Sokal et al. 2016) are consistent with the disruption of dense, star-forming gas on short (< 5 Myr) time-scales by the radiation and thermal pressure from young, massive stars (pre-supernova stellar feedback). Finally, the near-proportionality of the SFR surface density Σ_{SFR} and the molecular gas surface density Σ_{H_2} observed on kpc scales (e.g. Wong & Blitz 2002; Bigiel et al. 2008; Leroy et al. 2013) can also be explained by a self-regulation process, in which the density distribution of molecular gas is shaped by the stars it forms (Semenov, Kravtsov & Gnedin 2019).

However, the connection between this Lagrangian picture of star formation and the properties of observable star-forming regions or ‘giant molecular clouds’ remains unclear. Molecular clouds are observed to have a large range of masses, sizes, and densities, spanning over two orders of magnitude in Milky Way-like disc galaxies (e.g. Sun et al. 2018; Colombo et al. 2019). These star-forming regions are *not* universally destroyed on time-scales of 5–15 Myr. Both observations (e.g. Scoville & Hersch 1979; Engargiola et al. 2003; Blitz et al. 2007; Murray 2011; Corbelli et al. 2017)

* E-mail: sarah.jeffreson@cfa.harvard.edu

and simulations (e.g. Dobbs & Pringle 2013; Grisdale et al. 2019; Jeffreson et al. 2021a) show a large range of survival times for giant molecular clouds, ranging over two orders of magnitude, from 1 to >100 Myr. While short-lived (~ 10 Myr), low-mass ($\sim 10^4 M_{\odot}$) molecular clouds dominate by number, a substantial fraction of molecular mass (and thus galactic star formation) resides in the most massive ($\gtrsim 10^6 M_{\odot}$) molecular clouds (e.g. Murray 2011; Miville-Deschênes, Murray & Lee 2017; Faesi, Lada & Forbrich 2018), which tend to live for longer periods of time (e.g. Jeffreson et al. 2021a). As such, the mass-weighted molecular cloud lifetime is typically up to a factor of 10 longer than the star-forming lifetime of a Lagrangian gas parcel. We may therefore ask: How do we reconcile the short Lagrangian star formation time-scale with the long lifetimes of observable molecular clouds?

This question is closely related to another unanswered question in the field of star formation: What is the chemical lifetime of molecular hydrogen and how does it relate to the molecular cloud lifetime? Based on observations of massive molecular clouds in the low-density interarm regions of spiral galaxies, Scoville & Hersh (1979), Koda et al. (2009), and Koda, Scoville & Heyer (2016) argue that the survival time of hydrogen molecules must be comparable to the interarm crossing time. Given that the long time-scale for the formation of new hydrogen molecules from the low-density, interarm gas prohibits the formation of new massive clouds in these regions, these authors argue that high-mass molecular agglomerations must form within the spiral arms, then later fragment into smaller (but still massive) molecular clouds as they transit into the interarm regions. However, this hypothesis is at odds with the short Lagrangian lifetimes of star-forming gas (Semenov, Kravtsov & Gnedin 2017): when cold, dense, star-forming gas parcels are ejected into a warm, diffuse non-star-forming state by stellar feedback on time-scales of ~ 5 –15 Myr, their hydrogen molecules should also be dissociated by the interstellar radiation field.

The chemical lifetime of H_2 has important consequences for interpreting the observed spatial distribution of young star clusters and dense molecular gas, and in particular their decorrelation on sub-kiloparsec scales (e.g. Blitz et al. 2007; Kawamura et al. 2009; Schrubba et al. 2010; Miura et al. 2012; Corbelli et al. 2017; Kruijssen et al. 2019). If the H_2 survival time-scale is short, then the observed spatial decorrelation of young stars and H_2 on sub-kiloparsec scales implies that stellar feedback rapidly dissociates molecular gas and destroys the star-forming gas in its vicinity, providing evidence for short lifetimes of star-forming regions. However, if the H_2 survival time-scale is long, the same observations imply that the molecular gas is simply pushed away from young stars, where it can continue to form new stars.

In this work, we seek to reconcile the apparent contradiction between the Lagrangian and Eulerian views of molecular cloud lifetime, and to illuminate the relationship of cloud to chemical lifetimes. We use passive Lagrangian tracer particles in a chemodynamic simulation of an isolated disc galaxy in the moving-mesh code AREPO to directly compare the star-forming time-scale of Lagrangian gas parcels, the chemical lifetime of molecular hydrogen, and the lifetimes of giant molecular clouds. We describe our numerical prescription in Section 2, and examine the distribution of molecular mass and star formation among short- and long-lived molecular clouds in Section 3. We compare our molecular cloud lifetimes to the Lagrangian star formation time-scale and the molecular hydrogen survival time in Section 4, and use this result to develop a simple picture of molecular cloud evolution in Section 4.3. In Section 4.4, we also show that the massive, long-lived molecular clouds formed in this picture account for the clustering of supernova feedback in

our simulation, and are therefore the key drivers of galactic winds. We discuss our results in the context of the existing literature in Section 5, and present our conclusions in Section 6.

2 SIMULATION OF A DWARF SPIRAL GALAXY

Here, we describe our simulation setup and methods.

2.1 Initial conditions

We simulate a dwarf flocculent spiral galaxy that is analogous in its gas and stellar mass distribution to the bulgeless nearby galaxy NGC300. The initial condition includes a dark matter halo at a mass resolution of $1.254 \times 10^7 M_{\odot}$, a stellar disc at a mass resolution of $3437 M_{\odot}$, and a gas disc at a mass resolution of $859 M_{\odot}$. The dark matter halo follows the profile of Navarro, Frenk & White (1997), with a concentration parameter of $c = 15.4$, a spin parameter of $\lambda = 0.04$, a mass of $8.3 \times 10^{10} M_{\odot}$, and a circular velocity of $V_{200} = 76 \text{ km s}^{-1}$ at the virial radius. The stellar disc is of exponential form, with a mass of $1 \times 10^9 M_{\odot}$, a scale length of 1.39 kpc, and a scale height of 0.28 kpc. The corresponding exponential gas disc extends beyond the stellar disc, with a mass of $2.2 \times 10^9 M_{\odot}$ (giving a gas fraction of 68 per cent) and a scale length of 3.44 kpc.

2.2 Hydrodynamics and chemistry

The initial condition is evolved using the moving-mesh hydrodynamics code AREPO (Springel 2010). Within AREPO, the gas dynamics is modelled on the unstructured moving mesh defined by the Voronoi tessellation about a discrete set of points, which move according to the local gas velocity. The gravitational acceleration vectors of the Voronoi gas cells, stellar particles, and dark matter particles are computed using a hybrid TreePM gravity solver.

The temperature and chemical composition of the gas in our simulation is modelled using the simplified network of hydrogen, carbon and oxygen chemistry described in Nelson & Langer (1997) and in Glover & Mac Low (2007a, b), Smith et al. (2014). The fractional abundances of the chemical species H, H_2 , H^+ , He, C^+ , CO, O, and e^- are computed and tracked for each gas cell, and self-consistently coupled to the heating and cooling of the interstellar medium via the atomic and molecular cooling function of Glover et al. (2010). The gas equilibrates to a state of thermal balance between line-emission cooling and heating due to the photoelectric emission from polycyclic aromatic hydrocarbons and dust grains, as they interact with the interstellar radiation field (ISRF) and with cosmic rays. We assign a value of 1.7 Habing fields to the UV component of the ISRF (Mathis, Mezger & Panagia 1983), a value of $3 \times 10^{-17} \text{ s}^{-1}$ to the cosmic-ray ionization rate (van der Tak & van Dishoeck 2000), and we assume the solar value for the dust-to-gas ratio. We note that this results in a higher metallicity than the sub-solar value that is observed for NGC300 (Bresolin et al. 2009), but we demonstrate in Section 3.1 that our simulation nevertheless produces a realistic interstellar medium structure, with a realistic distribution of molecular cloud properties.

We use the TREECOL algorithm introduced by Clark, Glover & Klessen (2012) to model the dust- and self-shielding of molecular hydrogen from dissociation by the ISRF. This allows us to accurately model the non-equilibrium abundance of molecular hydrogen during the run-time of the simulation, and so to compute its value for each gas cell as a function of time.

2.3 Tracer particles

We introduce passive tracer particles to the simulation following the Monte Carlo prescription of Genel et al. (2013), which allows us to track the Lagrangian mass flow and molecular fraction of gas as it moves through simulated GMCs, despite the fact that AREPO is not a Lagrangian code. Via this prescription, tracer particles are moved along with the gas cells in the simulation, and are exchanged between gas cells according to a probability set by the mass flux between them. When a gas cell is converted to a star particle, the tracer particles associated with that gas cell are moved to the star particle with a probability set by the ratio of the stellar mass to the original gas cell mass. Similarly, tracer particles attached to star particles are transferred back to the gas reservoir according to the masses ejected in stellar winds and supernovae. As such, the mass of tracers in the gas and stellar reservoirs remains equal to the masses of these reservoirs throughout the simulation.

In this work, we assign one tracer particle to each gas cell in the initial condition, which sets an initial mass distribution for the tracer particles, equal to the initial mass distribution of gas cells. Since the cells are not of exactly equal mass initially, this means that the tracers are not uniform in mass either. However, we show in Appendix A that the distribution of effective tracer masses converges to a uniform value after after <100 Myr. We analyse the simulated disc only after it is in a state of dynamical equilibrium, between simulation times of 500 and 800 Myr. The effective tracer mass during this period is stable at a value of $\sim 450 M_\odot$, corresponding to 1.9 tracer particles per gas cell.

Finally, we note that the Monte Carlo nature of the tracer particles used in our simulation introduces a degree of non-physical spatial diffusivity, which cannot be improved upon by increasing their number relative to the gas cell resolution (see section 3.1 and fig. 6 of Genel et al. 2013). That is, the number of tracers does not change the probability that a given tracer is advected in a particular direction: this probability depends only on the mass flux between gas cells and the initial masses of those gas cells, as in their equations (1)–(5). Despite this, Monte Carlo tracers are still the better option relative to velocity field tracers, for grid-based codes (see second paragraph, section 6 of Genel et al. 2013).

2.4 Star formation and feedback

The star formation efficiency ϵ_{ff} per free-fall time of the gas in our simulations follows the parametrization of Padoan et al. (2017). These authors conduct high-resolution simulations of turbulent fragmentation and find that ϵ_{ff} depends on the local virial parameter α_{vir} of the gas, according to

$$\epsilon_{\text{ff}} = 0.4 \exp(-1.6\alpha_{\text{vir}}^{0.5}). \quad (1)$$

To use such a star formation recipe, simulations need a model for unresolved turbulent velocity dispersion, σ , to calculate $\alpha_{\text{vir}} \propto \sigma^2/(\rho L^2)$. To this end, existing methods in galaxy formation simulations include explicit dynamic models for sub-grid turbulence (e.g. Braun & Schmidt 2015; Semenov, Kravtsov & Gnedin 2016; Kretschmer & Teyssier 2020), estimation of σ from the gradients of resolved gas velocity (e.g. Hopkins, Narayanan & Murray 2013), or calculation of σ over some resolved scale (e.g. Gensior, Kruijssen & Keller 2020). As our simulations do not include a sub-grid model for turbulence, we follow the prescription of Gensior, Kruijssen & Keller (2020). In brief, the scale over which α_{vir} is calculated is computed using Sobolev approximation: $L = |\langle \rho_{\text{g}} \rangle / |\nabla \rho_{\text{g}}|$, where $\langle \nabla \rho_{\text{g}} \rangle$ is the cubic spline kernel-weighted average of the gas volume

density gradient, with respect to the radial distance from the central gas cell. The smoothing length of the cubic spline kernel is chosen to enclose the 32 nearest-neighbour cells. We refer the reader to Gensior, Kruijssen & Keller (2020) for a more detailed explanation. The SFR volume density of each gas cell in the simulation is therefore given by

$$\frac{d\rho_{*,i}}{dt} = \begin{cases} \frac{\epsilon_{\text{ff}} \rho_i}{t_{\text{ff},i}}, & \rho_i \geq \rho_{\text{thresh}} \text{ and } T_i \leq T_{\text{thresh}} \\ 0, & \rho_i < \rho_{\text{thresh}} \text{ or } T_i > T_{\text{thresh}} \end{cases}, \quad (2)$$

where $t_{\text{ff},i} = \sqrt{3\pi/(32G\rho_i)}$ is the local free-fall time-scale for the gas cell i with a mass volume density of ρ_i , and ϵ_{ff} is given by equation (1).

We set a lower limit of $\rho_{\text{thresh}}/m_{\text{H}}\mu = 100 \text{ cm}^{-3}$ on the volume density of hydrogen atoms above which star formation is allowed to occur, as well as an upper limit of $T_{\text{thresh}} = 100 \text{ K}$ on the temperature; here $\mu \approx 1.4$ is the mean mass per H atom. The value of ρ_{thresh} is the density of Jeans-unstable (collapsing) gas at our mass resolution of $859 M_\odot$ and at the temperature of $\sim 30 \text{ K}$ reached by the molecular gas in our simulation. For a spherical gas cell, the radius associated with this density threshold is 3 pc, and so we employ the adaptive gravitational softening scheme in AREPO with a minimum softening length of 3 pc and a gradation of 1.5 times the Voronoi gas cell diameter. We set the softening length of the stellar particles to the same value, and choose a softening length of 280 pc for the dark matter particles, according to the convergence tests presented in Power et al. (2003). Because our simulations resolve the gas-disc scale height and the Toomre mass at all scales, the adaptive gravitational softening avoids the majority of artificial fragmentation at scales larger than the Jeans length (Nelson 2006), as discussed in greater detail in Jeffreson et al. (2020) and in Appendix C.

The stellar feedback in our simulation is modelled via injection of energy and momentum by supernovae and pre-supernova H II regions. To compute the number of supernovae generated by each star particle formed during the simulation, we assign a stellar population drawn stochastically from a Chabrier (2003) initial stellar mass function, using the Stochastically Lighting Up Galaxies (SLUG) stellar population synthesis model (da Silva, Fumagalli & Krumholz 2012, 2014; Krumholz et al. 2015). Within SLUG, the resulting stellar populations are evolved along Padova solar metallicity tracks (Fagotto et al. 1994a, b; Vázquez & Leitherer 2005) during run-time, using STARBURST99-like spectral synthesis (Leitherer et al. 1999). This modelling provides the number of supernovae $N_{*,\text{SN}}$ generated by each star particle during each simulation time-step, as well as the ionizing luminosity of the cluster and the mass Δm_* it has ejected.

We use the values of $N_{*,\text{SN}}$ and Δm_* for each star particle to compute the momentum and thermal energy injected by supernova explosions at each time-step. In the case of $N_{*,\text{SN}} = 0$, we assume that all mass-loss results from stellar winds. In the case of $N_{*,\text{SN}} > 0$, we assume that all mass-loss results from supernova explosions, and we model the corresponding kinetic and thermal energy injection due to the expanding blast-wave. At our mass resolution of $859 M_\odot$ per gas cell, the energy-conserving/momentum-generating phase of supernova blast-wave expansion is unresolved, and so we follow the prescription introduced by Kimm & Cen (2014): we explicitly inject the terminal momentum of the blast-wave into the set of gas cells k that share faces with the nearest-neighbour cell to the star particle. We use the unclustered parametrization of the terminal momentum derived from the high-resolution simulations of Gentry et al. (2017), which is given by

$$\frac{P_{\text{t},k}}{M_\odot \text{ km s}^{-1}} = 4.249 \times 10^5 N_{*,\text{SN}} \left(\frac{n_k}{\text{cm}^{-3}} \right)^{-0.06}, \quad (3)$$

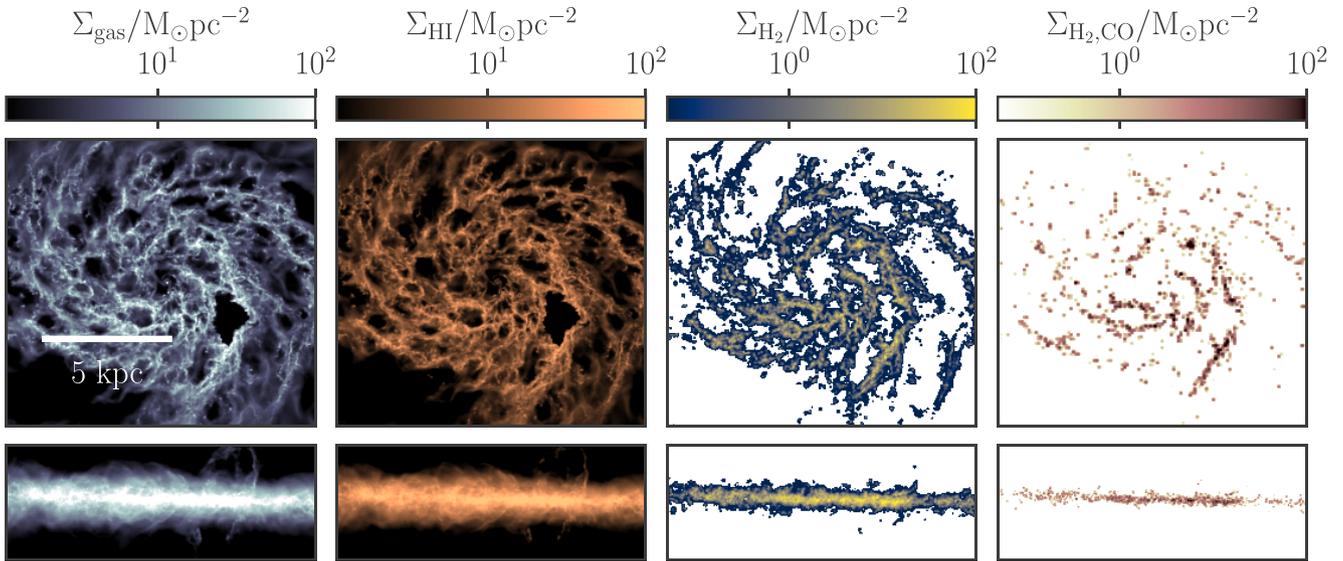


Figure 1. Column density maps of the total (Σ_{gas} , left-hand panel), atomic (Σ_{HI} , centre left-hand panel), total molecular (Σ_{H_2} , centre right-hand panel), and CO-luminous molecular ($\Sigma_{\text{H}_2,\text{CO}}$ right-hand panel) gas distribution for the simulated dwarf spiral galaxy, viewed perpendicular to (top panels) and across (lower panels) the galactic mid-plane, at a simulation time of 800 Myr.

with an upper limit imposed by the condition of kinetic energy conservation, as the shell sweeps through the gas cells k . The momentum is distributed among the facing cells as described in Jeffreson et al. (2021b).

The pre-supernova feedback from H II regions is implemented according to the model of Jeffreson et al. (2021b). This model accounts for the momentum injected by both radiation pressure and the ‘rocket effect’, whereby warm ionized gas is ejected from cold molecular clouds, following the analytic work of Matzner (2002) and Krumholz & Matzner (2009). Momentum is injected for groups of star particles with overlapping ionization-front radii, which are identified using a Friends-of-Friends grouping prescription. The momentum is received by the gas cell closest to the luminosity-weighted centre of the Friends-of-Friends group, and is distributed to the set of adjoining neighbour cells. The gas cells inside the Strömgren radii of each Friends-of-Friends group are heated to a temperature of 7000 K, and are held above this temperature floor for as long as they receive ionizing photons from the group. In contrast to Jeffreson et al. (2021b), we also explicitly and fully ionize the gas inside these Strömgren radii, rather than relying on the chemical network to do so.

3 PROPERTIES OF THE SIMULATED GALAXY

Here, we compute the key observable properties of our simulated galaxy and its interstellar medium, and demonstrate that their values are close to those obtained via direct observations in comparable galactic environments.

3.1 Gas morphology and global interstellar medium structure

Fig. 1 shows the spatial distribution of the total (left-hand panel), atomic (centre left-hand panel), total molecular (centre right-hand panel) and CO-luminous molecular (right-hand panel) gas reservoirs at face-on and edge-on viewing angles, at a simulation time of 800 Myr. The atomic and total molecular gas abundances are

computed within our simplified chemical network during simulation run-time, as described in Section 2. To partition the molecular gas into its CO-luminous and CO-dark components, we post-process our simulation outputs using the DESPOTIC astrochemistry and radiative transfer model (Krumholz 2013, 2014), as described in Appendix B, and delineate CO-dark from CO-bright material using a threshold that we detail below.

The qualitative structure of the CO-bright interstellar medium is similar to that observed by Faesi, Lada & Forbrich (2018) in the nearby dwarf spiral galaxy NGC300, displayed in fig. 1 of Kruijssen et al. (2019). In Fig. 2, we quantify the agreement between the structure of our simulated interstellar medium and the observed interstellar medium in NGC300, showing that the simulation reproduces the observed spatial decorrelation between regions of recent star formation (traced by young stars) and regions of high molecular gas surface density (traced by CO emission). To this end, we use the so-called ‘tuning fork diagram’ that quantifies the relative bias of depletion time measured in the apertures centred either on dense gas or young stars (e.g. Schruba et al. 2010; Kruijssen et al. 2019). The relative excess of dense gas or young stars in apertures of a given size (x -axis) results in the upper (blue) and lower (green) branches, respectively. The values calculated from observations by Kruijssen et al. (2019) are given by the thin lines and open data points.

We calculate the dense gas and young stellar branches for our simulation as outlined in Semenov, Kravtsov & Gnedin (2021). In particular, we use molecular gas with a surface density of $\Sigma_{\text{H}_2,\text{CO}} > \Sigma_{\text{H}_2,\text{CO},\text{min}} = 13 M_{\odot} \text{pc}^{-2}$ and a line-of-sight velocity dispersion of $< 1 \text{ km s}^{-1} \times \Sigma_{\text{H}_2,\text{CO}} / \Sigma_{\text{H}_2,\text{CO},\text{min}}$ as a proxy for the dense CO-bright gas; here, $\Sigma_{\text{H}_2,\text{CO}}$ is the H_2 surface density that would be inferred based on the computed CO luminosity, rather than the true H_2 surface density. We define this quantity precisely in Appendix B. The line-of-sight velocity dispersion is computed as the CO-weighted mean velocity dispersion of the Voronoi gas cells along rays perpendicular to the galactic mid-plane. We assume that the $\text{H}\alpha$ signal in young star-forming regions is produced by star particles with ages 2–5 Myr.

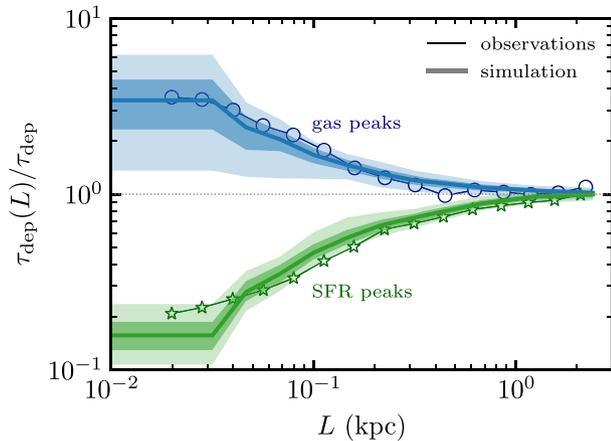


Figure 2. Our simulated dwarf spiral galaxy reproduces the spatial decorrelation observed for NGC300 between young stars and dense gas on the scales of molecular clouds – the so-called ‘tuning fork diagram’ (Kruijssen et al. 2019, thin lines with markers). The branches of the tuning fork show the average bias of gas depletion times measured in the apertures of a given size (x -axis) placed on the peaks of dense gas (top branch) or young stars (lower branch). To quantify the snapshot-to-snapshot variation of the tuning fork in our simulation, the thick solid lines show medians and the shaded regions show 16–84 and 2.5–97.5 inter percentile ranges. The diagram is computed over the range of galactocentric radii of $R = 2\text{--}6$ kpc (see the text for details).

The shaded regions show the 1σ and 2σ spread of values for our simulated interstellar medium between simulation times of 500 and 800 Myr.

The convergence, on large scales, of the upper and lower branches in Fig. 2 is a general property of star-forming galaxies (e.g. Kim et al. 2022), indicating the tight correlation between dense gas and young stars that is manifested in the near-linear relation between the molecular gas density and the star formation density observed in normal (non-starburst) star-forming galaxies on $>$ kiloparsec scales (e.g. Wong & Blitz 2002; Bigiel et al. 2008; Leroy et al. 2013). On small scales, the branches diverge, as dense, CO-emitting gas and ionized, $H\alpha$ -emitting gas become spatially decorrelated.

The precise shape of the branches in Fig. 2 is sensitive to the modelling of star formation and stellar feedback used in the numerical simulation (Fujimoto et al. 2019; Semenov, Kravtsov & Gnedin 2021). The agreement of this shape with the relation computed from observational data by Kruijssen et al. (2019) indicates that our models produce a reasonable interstellar medium structure for galactocentric radii of $R = 2\text{--}6$ kpc. The inner 2 kpc of the simulation displays a stronger correlation (narrower branch opening) than is computed by Kruijssen et al. (2019). Although this is likely due to the existence of several high-mass molecular clouds in this region, which are removed (‘masked’) in their analysis, we exercise caution and exclude the region $R < 2$ kpc from our subsequent analysis, as we cannot verify its agreement with the observational data.

For the remaining galactocentric radii $R > 2$ kpc, we demonstrate in Fig. 3 that our simulated dwarf spiral galaxy is similar in its atomic and molecular depletion times (upper panel) and its gas and stellar surface densities (lower panel) to NGC300. The simulated values are given by the bold lines, while the corresponding observed values from Westmeier, Braun & Koribalski (2011) and Kruijssen et al. (2019) are given by the thin lines.

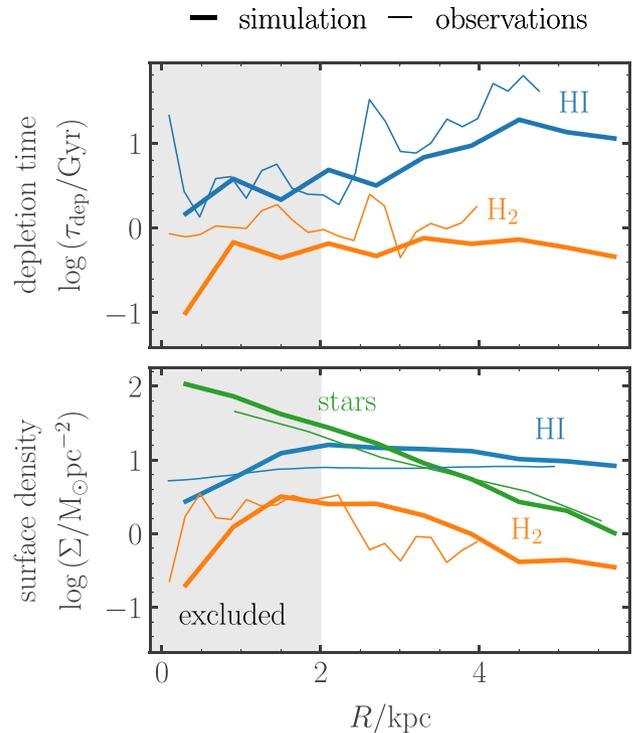


Figure 3. Our simulated dwarf spiral galaxy is similar to NGC300 in its gas and stellar surface densities and gas depletion times. Upper panel: molecular and atomic gas depletion times as a function of galactocentric radius (solid lines), measured within annuli of width 500 pc. Thin lines show the values measured for NGC300 by Kruijssen et al. (2019). Lower panel: stellar, molecular, and atomic gas surface densities as a function of galactocentric radius, measured within the same annuli. Thin lines show the values measured for NGC300 by Kruijssen et al. (2019) and by Westmeier, Braun & Koribalski (2011).

3.2 The CO-bright molecular cloud population

3.2.1 Identification

In this work, we identify CO-bright, observable clouds using isocontours of value $\log(\Sigma_{H_2,CO}/M_{\odot}pc^{-2}) = -1.5$ on the surface density $\Sigma_{H_2,CO}$ of CO-bright molecular hydrogen perpendicular to the galactic mid-plane, shown in dark pink in Fig. 4. This threshold corresponds to the natural break in the distribution of $\Sigma_{H_2,CO}$ produced by our chemical post-processing, which is described in detail in Appendix B. At surface densities higher than the threshold, gas cells contain at least some shielded, CO-dominated gas. At lower densities, CO exists only as a uniformly mixed, unshielded, low-abundance component.

The yellow contours in Fig. 4 enclose regions with a projected H_2 abundance of $\Sigma_{H_2}/\Sigma_{gas} > 0.3$, accounting for 70 per cent of the total H_2 in the simulation, including CO-dark H_2 . We find that the majority of star formation (78 per cent) occurs in the CO-luminous H_2 , despite the fact that this gas reservoir accounts for only 26 per cent of the galactic molecular hydrogen mass. In what follows, we will therefore focus on the CO-luminous giant molecular clouds. We note that 99 per cent of the gas tracer particles occupying the CO-luminous state inside the pink contours of Fig. 4 have previously resided in the CO-dark, H_2 -rich state denoted by the yellow contours. That is, the gas in our simulation almost always enters CO-luminous clouds by way of the CO-dark envelope.

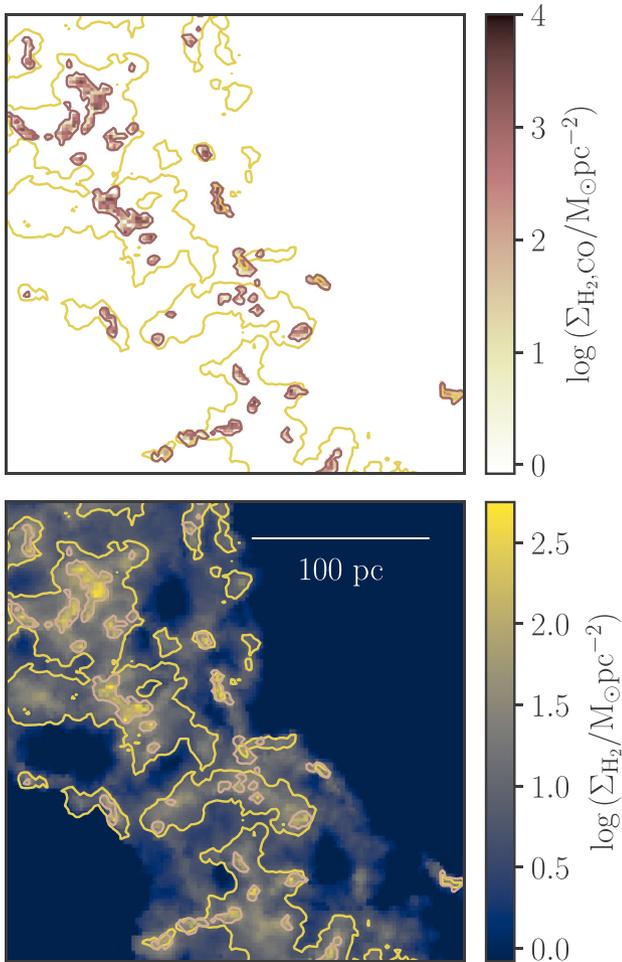


Figure 4. Thresholds for the identification of CO-luminous giant molecular clouds (pink isocontours) and their CO-dark envelopes down to a fractional projected abundance of 0.3 (yellow isocontours). The upper panel shows the CO-luminous molecular hydrogen column density (pink) and the lower panel shows the total molecular hydrogen column density (blue–yellow).

3.2.2 Observational checks

An important check for the applicability of our simulation to the real Universe is whether the simulated population of CO-bright molecular clouds has similar properties to the observed population in a dwarf spiral galaxy. In Fig. 5, we compare the instantaneous mass distribution (left-hand panel) and size distribution (centre panel) of our simulated molecular clouds across all simulation times to the population of resolved clouds in the inner disc ($R \lesssim 3$ kpc) of NGC300 (Faesi, Lada & Forbrich 2018) and in the outer regions ($R \gtrsim 2$ kpc) of M33 (Gratier et al. 2012). We divide the simulated cloud population into four mass bins, to highlight the correspondence between the cloud mass, surface density and velocity dispersion.

We see that the slope of the simulated cloud mass distribution agrees well with the observed slope of $\beta = 2 \pm 0.1$ in the outer regions of M33, but is significantly steeper than the corresponding slope in NGC300. It does not display a truncation at high masses, as seen for NGC300 (not shown in the figure). Given that our simulated cloud population is made up of clouds at $R > 2$ kpc, we would expect that our results are more closely comparable to the M33 sample. Note, however, that such a comparison should be interpreted with caution, given the sensitivity of the mass function slope to the method of cloud

identification used (Pineda, Rosolowsky & Goodman 2009; Hughes et al. 2013; Rosolowsky et al. 2021). Nevertheless, we note that the distribution of masses for the molecular clouds in our simulation is in reasonable agreement with observed values.

Comparing the coloured contours to the coloured lines in the right-hand panel, we see that the cloud population generally follows a line of constant virial parameter, as seen in the observations of resolved clouds in NGC300 (black data points, Faesi, Lada & Forbrich 2018). Higher mass clouds are closer to virial equilibrium on average, with higher levels of gravitational boundedness.

3.2.3 Tracking and evolution

We track the temporal evolution of the simulated molecular cloud population via the algorithm described in Jeffreson et al. (2021a). In brief, the position of the two-dimensional isocontour enclosing a star-forming region at simulation time t is projected forward by the time-step $\Delta t = 1$ Myr of our simulation output, using the velocities of the gas cells in the region. A pair of star-forming regions is temporally linked as parent and child if there exists any overlap between the projected contour of the parent at time t and the contour outlining the child at time $t + \Delta t$.

Via this tracking procedure, we produce the *cloud evolution network* for the entire simulation, composed of ~ 8000 complete, independent segments between simulation times of $t = 500$ and 800 Myr, and between galactocentric radii of $R = 2$ and 6 kpc. These segments correspond to time-evolving molecular clouds, which we use in the analysis presented in Section 4. The lifetimes of these clouds are calculated as the end-to-end time between the formation of the first parent cloud and the destruction of the last child, accounting for all cloud mergers and splits, as shown in Fig. 6.

4 RESULTS

In this section, we combine our analysis of the molecular cloud network with results derived from the tracer particles in our simulation to study the formation and destruction of H_2 molecules within CO-bright molecular clouds, as a function of the cloud mass and cloud lifetime. We begin with the distribution of masses and lifetimes derived from the tracking of Eulerian clouds in Section 4.1, turn to the results from the tracer particles in Section 4.2, and combine the two viewpoints in Section 4.3. In Section 4.4, we discuss the impact of the Eulerian cloud mass distribution on the clustering of supernovae.

4.1 The Eulerian view: long-lived clouds

In the upper panel of Fig. 7, we show the cumulative probability distribution of CO-bright molecular clouds in our simulation, as a function of their lifetimes. In the second panel, we show the corresponding cumulative fraction of the galactic H_2 reservoir (orange) and SFR (purple) that is accounted for by these clouds. We see that half of the star formation occurring in CO-bright molecular clouds (40 per cent of the galactic total) is accounted for by molecular clouds that live longer than 25 Myr, despite the fact that the most-common lifetime for molecular clouds, by number, is ~ 3 Myr. Clouds with lifetimes > 25 Myr contain half the total CO-luminous H_2 mass in the galaxy.

In the third panel of Fig. 7, we demonstrate that molecular cloud lifetime is tightly correlated and monotonically increasing with the peak cloud mass achieved throughout this lifetime. We will

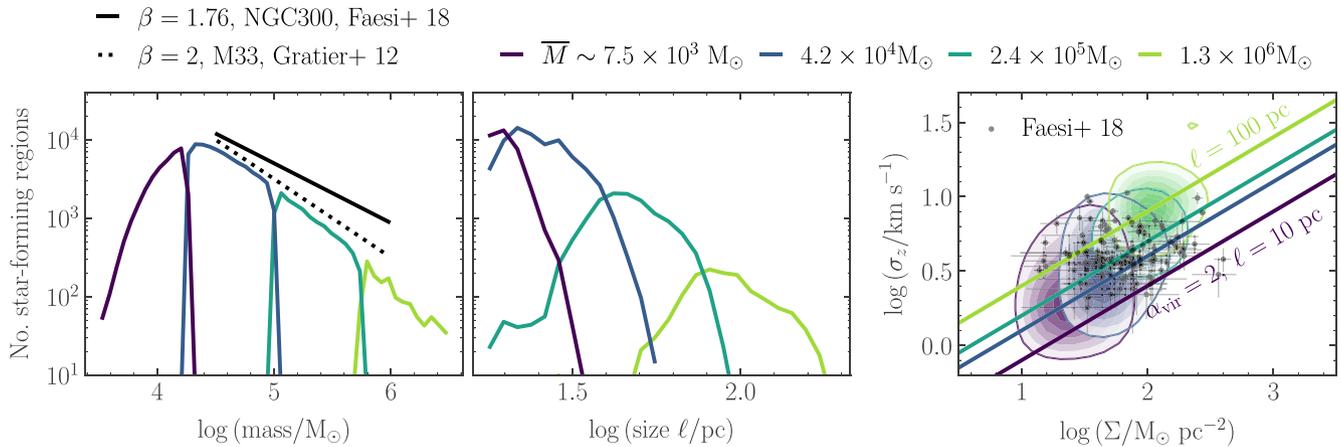


Figure 5. Left-hand panel: mass distribution of CO-bright molecular clouds in our simulation, divided into four mass bins. The solid black line gives the observed power-law slope $dN/dM \propto M^{-\beta}$, $\beta = 1.76 \pm 0.07$ found by Faesi, Lada & Forbrich (2018) in NGC300. The dotted black line gives the power-law slope $\beta = 2 \pm 0.1$ found by Gratier et al. (2012) in M33. Centre panel: size distribution of CO-bright molecular clouds in our simulation, divided into the same four mass bins. Right-hand panel: molecular gas line-of-sight velocity dispersion as a function of the molecular gas surface density for the CO-luminous molecular clouds. The solid lines indicate virial parameters of $\alpha_{\text{vir}} = 2$ for spherical beam-filling clouds at the mean region size for each mass bin. Black data points and error bars represent the resolved GMC sample of Faesi, Lada & Forbrich (2018) in NGC300.

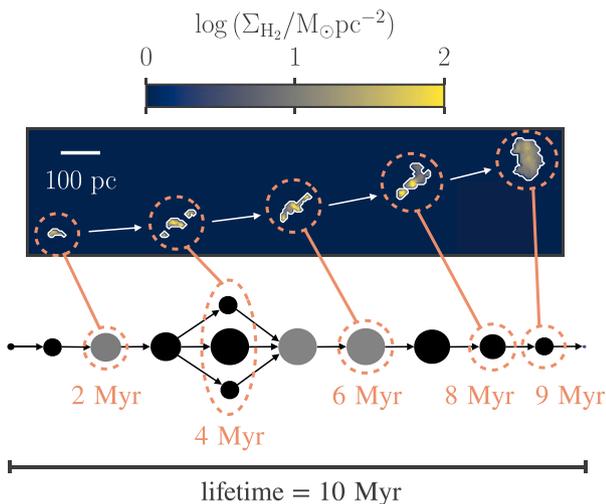


Figure 6. The lifecycle of a single simulated molecular cloud with a lifetime of 10 Myr. The actual evolution of the H_2 surface density as it moves across the galaxy is shown in the upper panel. Both the CO-bright and CO-dark H_2 are shown. This region undergoes a split and re-merger 4 Myr after its birth. Black filled circles correspond to time-steps during which the GMC forms no stars ($\text{SFR} = 0$), and grey filled circles to times during which the GMC forms stars ($\text{SFR} > 0$). The corresponding section of the cloud evolution network is shown in the lower half of the figure.

discuss this proportionality in detail in Section 4.3, but for now we simply note that ‘long-lived’ molecular clouds are synonymous with ‘massive’ clouds. Therefore, a substantial fraction of the CO-bright molecular gas in our dwarf spiral galaxy simulation is contained in clouds that are both massive and long-lived.

Finally, in the lower panel of Fig. 7, we show the number of tracer particles that transit through each simulated molecular cloud. These are the populations of tracer particles that we will analyse in Sections 4.2 and 4.3. We see that even in the lowest mass, shortest-lived GMCs studied, the median number of tracer particles per

cloud is ~ 100 . Additionally, we study only the average behaviour of large groups of GMCs in this work, such that the minimum tracer population size analysed is 3000 (see Fig. 9). This is a sufficient number to ensure that the average behaviour of the tracer particles analysed is not dominated by the Poisson noise associated with their probabilistic advection.

4.2 The Lagrangian view: short-lived H_2 molecules

We now examine the chemical survival time of molecular hydrogen in CO-bright molecular clouds, by tracking the H_2 abundance of Lagrangian gas parcels (tracer particles) as they transit through these clouds. We consider a gas parcel to have passed through a cloud if it moves within the isocontour defining its edge at any time. We show the time-evolution of the H_2 mass fraction $x_{H_2} = \rho_{H_2}/\rho_{\text{gas}}$ for the collection of gas parcels transiting through a single example cloud in Fig. 8. We omit tracer particles that remain at molecular hydrogen abundances below $x_{H_2} = 10^{-5}$ throughout the simulation run-time from our analysis; this cut excludes those tracer particles that are far from the galactic mid-plane but appear inside the isocontour in projection.

As the plot shows, individual fluid elements experience a wide range of variations in x_{H_2} , with many experiencing rapid fluctuations in the H_2 abundance, particularly as the cloud is disrupted. To quantify this behaviour, for each tracer particle that passes through a CO-bright molecular cloud, we identify all the contiguous time periods for which it has $x_{H_2} < 0.5$ and > 0.5 . We record the duration of each time period, the median (in time) value of x_{H_2} during each time period, and the lifetime of the CO-bright cloud through which the tracer particle has passed.¹ In the left- and right-hand panels of

¹For particles that have multiple episodes of $x_{H_2} > 0.5$ with a period of $x_{H_2} < 0.5$ between them, we take the molecular cloud lifetime for the H_2 -poor phase to be the lifetime of the molecular cloud through which the particle passed during the preceding H_2 -rich phase. Essentially no tracer particles pass through two different CO-bright molecular clouds during a single cycle of having $x_{H_2} > 0.5$ and < 0.5 .

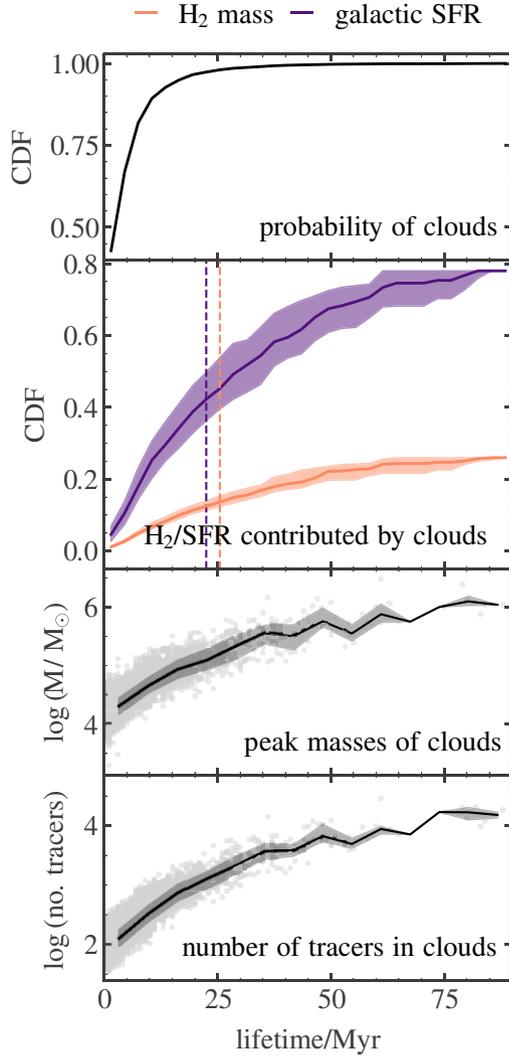


Figure 7. Upper panel: cumulative probability distribution of the molecular cloud lifetime by number. Second panel: time-averaged cumulative fraction of the total galactic molecular mass (orange) and SFR (purple) accounted for by CO-bright clouds, as a function of their lifetime. The solid lines denote the time-averaged median values, while the shaded regions denote the corresponding interquartile ranges. The normalization of the CDFs reflects the fact that CO-dark gas accounts for 22 per cent of star formation and 74 per cent of galactic H_2 (see Section 3.2.1). The orange and purple dashed vertical lines represent the corresponding weighted median cloud lifetimes (clouds above/below this lifetime account for 50 per cent of the H_2 mass or galactic SFR, respectively). Third panel: Median and interquartile range of the GMC peak mass, as a function of the cloud lifetime. Grey circles represent the values for individual GMCs. Lower panel: median and interquartile range of the total number of tracer particles that transits through each GMC, as a function of the cloud lifetime.

Fig. 9, we plot the median x_{H_2} versus duration, with points colour-coded by molecular cloud lifetimes; the left-panel shows intervals of $x_{\text{H}_2} > 0.5$, while the right shows $x_{\text{H}_2} < 0.5$. The plot shows that gas rapidly cycles in and out of an H_2 -dominated state ($x_{\text{H}_2} > 0.5$), in qualitative agreement with Semenov, Kravtsov & Gnedin (2017). The time spent by gas parcels in the H_2 -rich state is short, ranging from 1 to ~ 20 Myr with a median value of 4 Myr. The median abundance of molecular hydrogen over the duration of each cycle is 0.63. Crucially, the distributions of H_2 molecule lifetimes and x_{H_2} values are almost identical between the samples of molecular clouds

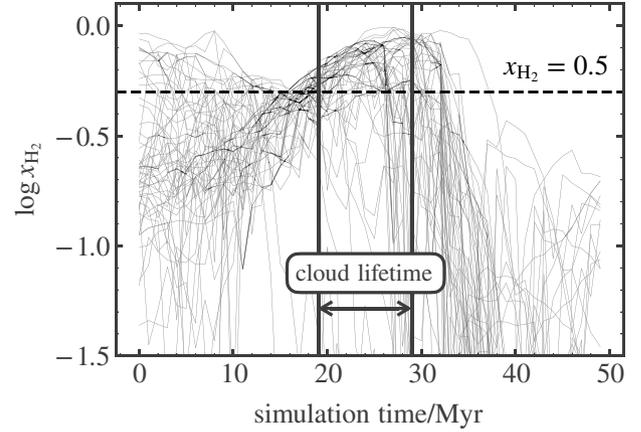


Figure 8. The evolution of the molecular hydrogen abundance x_{H_2} of individual Lagrangian gas parcels (thin black lines), as they transit through a CO-bright molecular cloud of lifetime 10 Myr. This is the same molecular cloud shown in Figure 6. The black vertical lines denote the times at which the cloud appears and disappears, based on our cloud-tracking algorithm.

with different lifetimes (different colours). Thus, we conclude that independent of the lifetime of the host molecular cloud, the chemical survival time of H_2 molecules is about 4 Myr, subject to the accuracy of our chemical model, which we discuss in detail in Section 5.2.

The right-hand panel of Fig. 9 shows that the time spent by gas parcels in the H_2 -poor state ($x_{\text{H}_2} < 0.5$) is correlated with the lifetime of the host molecular cloud, albeit weakly. Gas parcels ejected from higher mass, longer lived clouds (yellow) return to the H_2 -dominated state after shorter periods of time, and retain higher median H_2 abundances, than those associated with shorter lived, lower mass clouds (purple). That is, H_2 is not as efficiently or quickly dissociated in higher mass molecular clouds. Gas parcels ejected from these clouds are more likely to become ‘trapped’, and cycle back to an H_2 -dominated state multiple times: the typical number of cycles for tracer particles transiting through clouds of lifetime < 25 Myr is 1, but for regions of lifetime > 25 Myr, the number of cycles is between 2 and 4.

In Fig. 10, we compare the entire (unweighted) distribution of GMC lifetimes (purple, left-hand panel) to the entire distribution of H_2 -dominated periods for the tracer particles passing through these GMCs (orange, left-hand panel). The overall median values of each distribution are given by the thin-dashed lines, and the corresponding median lifetimes and time-scales, as a function of the peak mass reached during each GMC lifetime, are shown in the right-hand panel. We see that for GMCs with lifetimes ≤ 10 Myr, the cloud lifetime distribution closely matches the chemical survival time of H_2 : The GMCs are dispersed during the first period of H_2 destruction. These low-mass, short-lived clouds dominate the GMC population by number, as indicated by the median cloud lifetime of ~ 3 Myr (dashed purple line). By contrast, GMCs that reach masses of $\gtrsim 6 \times 10^4 M_{\odot}$ survive longer than their constituent H_2 molecules: up to ~ 90 Myr, relative to a median H_2 destruction time of 4 Myr. These higher mass, longer lived clouds dominate the GMC population by mass, accounting for 65 per cent of the molecular mass of the cloud population (recall Fig. 7).

In Fig. 11, we demonstrate that the ejection of gas from the H_2 -dominated state is driven by early stellar feedback. To construct this figure, we tag every Lagrangian tracer particle that lies within 10 pc of a star particle at the moment of its formation. For each tagged tracer, we also identify the lifetime of the GMC within which that tracer resides. On the left-hand side of Fig. 11, we plot the time-evolution of

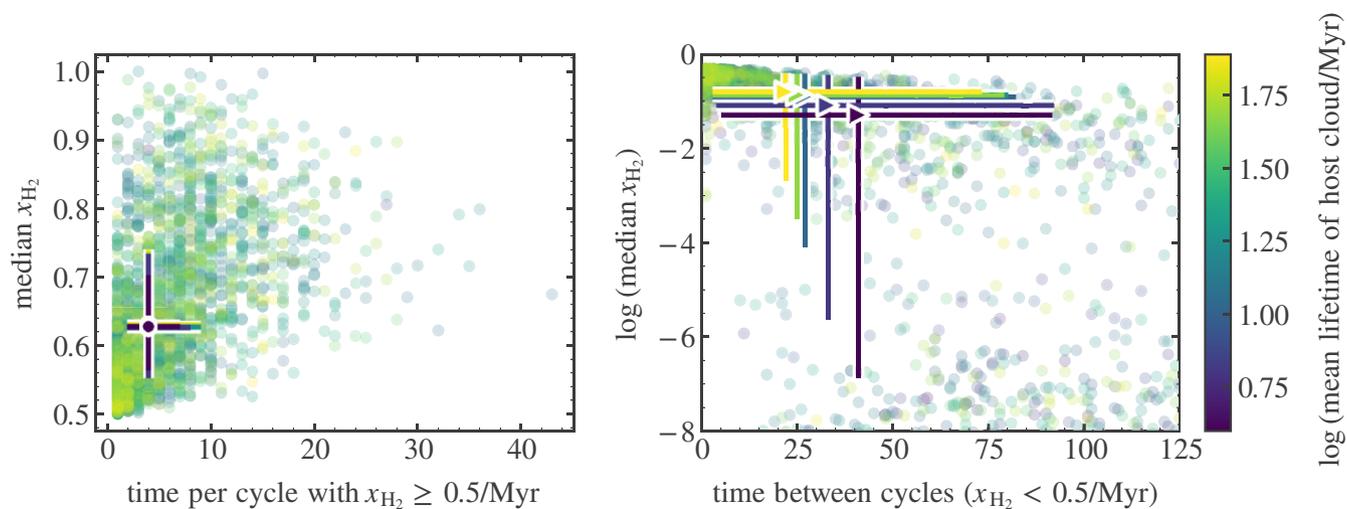


Figure 9. The time spent by gas tracer particles in the H_2 -rich ($x_{H_2} \geq 0.5$; left-hand panel) and H_2 -poor ($x_{H_2} < 0.5$; right-hand panel) phases, for tracer particles that transit through CO-bright molecular clouds. The transparent data points give the values for 1/5000th of the cycles undergone by tracer particles (each data point represents a tracer particle), while the solid data points and lines give the median values and interquartile ranges along each axis, computed in bins of molecular cloud lifetime as indicated by the colours of the solid points. All points are coloured according to the lifetime of the cloud through which the tracer particle transits. The time-scales spent in the H_2 -poor state are strictly lower limits, as 60 per cent of the values at the beginning or end of the simulation window at $t = 500$ or 800 Myr, and so are truncated. The smallest tracer particle population averaged in this figure (lowest cloud lifetime, in purple) is 3.0×10^3 tracer particles.

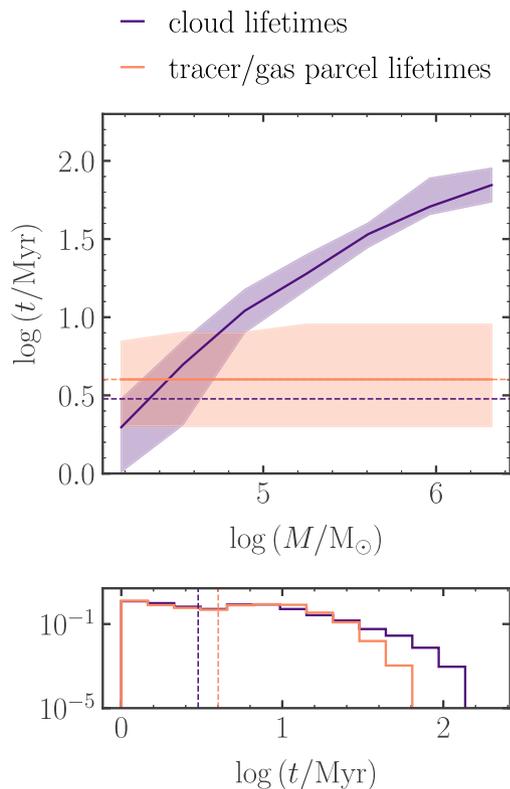


Figure 10. Lower panel: the entire normalized distribution of simulated GMC lifetimes (purple), relative to the entire normalized distribution of time intervals spent by gas tracer particles in the H_2 -rich phase (orange), as they transit through these GMCs. The median of each distribution is given by the thin dashed lines. Top panel: the corresponding median GMC lifetimes and tracer time-scales, plotted as a function of the peak cloud mass.

the H_2 fraction, total gas density, and distance from the star for each of these tracers, terminating each line at the time of the star particle’s first supernova. The thin lines in the figure correspond to the histories of individual tracers, and the thick lines to median values we obtain by binning the tracers by molecular cloud lifetime. Although the variation of individual tracks is large, we see that on average gas is efficiently ejected from the H_2 -dominated state by pre-supernova feedback on time-scales between 2 and 5 Myr. The gas is pushed out to a distance of ~ 80 pc and down to a density of between 1 and 10 cm^{-3} (close to the bulk mean density of hydrogen in the interstellar medium), where its H_2 is dissociated by the interstellar radiation field. However, the efficiency of this ejection is a strong function of molecular cloud lifetime: gas parcels in higher mass, longer lived molecular clouds start at higher densities and H_2 fractions, and after star formation they retain higher density and H_2 abundance out to larger distances from the newly formed star particle.

On the right-hand side of Fig. 11, we show the further time-evolution of these same Lagrangian tracers, but starting at the instant of the first supernova. The plot demonstrates that the role of supernova feedback in destroying H_2 is secondary to that of pre-supernova feedback. By the time that stars explode as supernovae, the surrounding gas is already relatively H_2 -poor ($x_{H_2} < 0.5$) and close to the mean density of the bulk interstellar medium ($\sim 1 \text{ cm}^{-3}$). The main (partial) exception to this is in the most massive and longest-lived clouds, where supernovae do further decrease the H_2 fraction. However, even for these clouds, the mean H_2 fraction is below 50 per cent by the time the first supernovae occur.

The increased difficulty of destroying H_2 molecules in higher mass/long-lived molecular clouds by pre-supernova feedback results in an increase in the integrated star formation efficiency with cloud mass and lifetime. In Fig. 12, we calculate the efficiency as the fraction of tracer particles that are transferred from the gas to star particles as they transit through molecular clouds, and find an increase from ~ 1 to ~ 4 per cent for an increase in cloud lifetime from ~ 10 to ~ 90 Myr. The horizontal dashed line represents the median integrated star formation efficiency for the entire CO-luminous gas reservoir. This

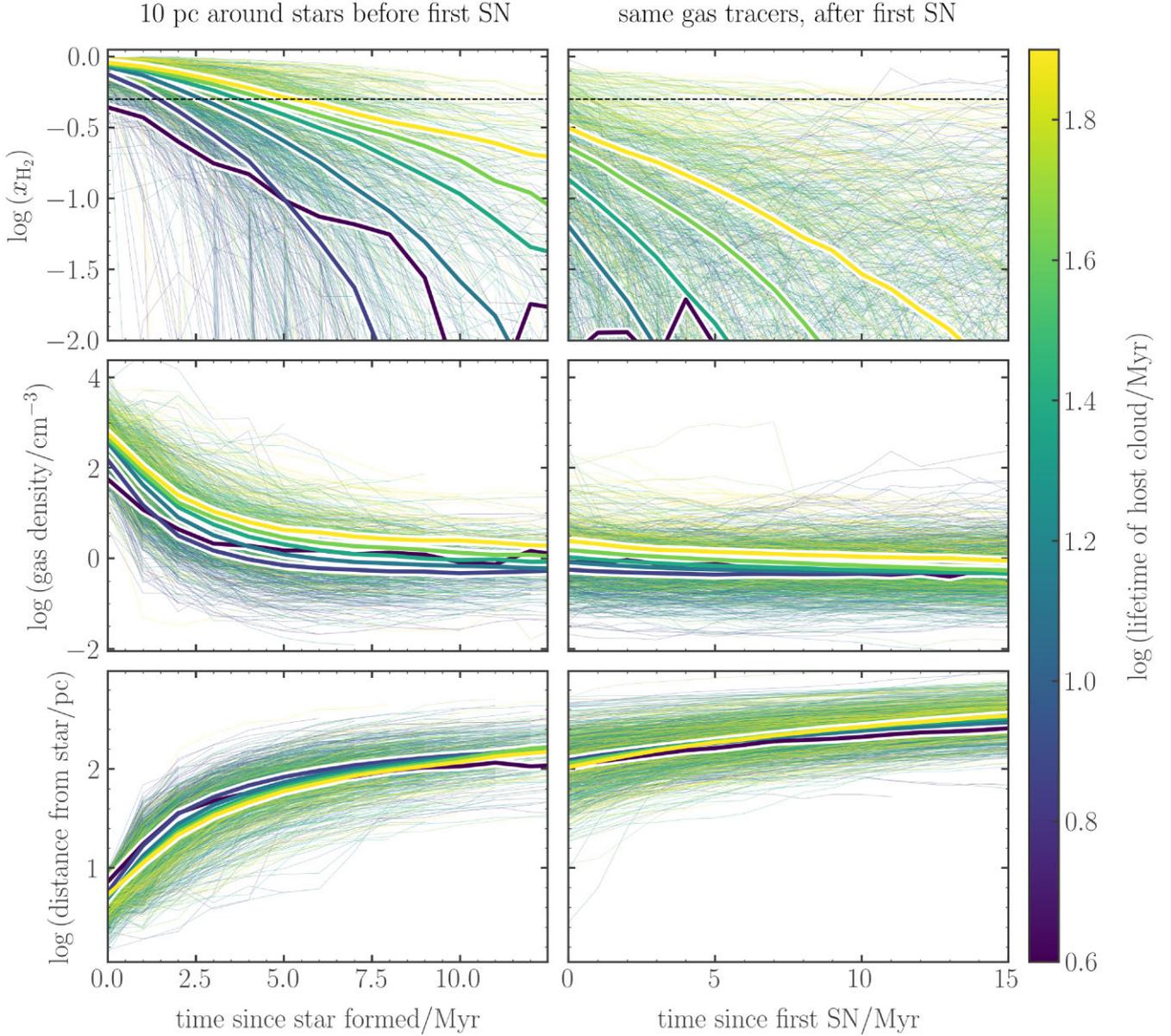


Figure 11. The behaviour of gas parcels (tracer particles) around all young star particles in our simulation. The colourbar corresponds to the lifetime of the host molecular cloud. Thin lines show the median values for the gas tracer particles that are located within 10 pc of a star particle at the instant of its formation, while thick lines show the medians across this population, computed in bins of molecular cloud lifetime. The left-hand column shows the time evolution starting at the instant of star formation, while the right-hand column shows the time evolution starting from the moment when the star particle produces its first supernova. We see that early, pre-supernova feedback (left-hand column) ionizes and drives the gas away from young stars on time-scales between 2 and 5 Myr across all molecular clouds, with the ejection being slightly slower in the longer lived, higher mass regions. Supernovae (right-hand column) generally occur when the gas has already been expelled from an H_2 -rich state. The smallest tracer particle population averaged in this figure (lowest cloud lifetime, in purple) is 1.7×10^4 tracer particles.

corresponds to the star formation efficiency associated with the most common cloud lifetime of ~ 10 Myr, as reported in Fig. 7.

Interestingly, Fig. 12 shows that the lifetimes of molecular regions, t_{GMC} , correlate near-linearly with their integrated SFE, ϵ_{int} . Such a correlation might explain the near-linear correlation between Σ_{H_2} and $\dot{\Sigma}_*$ observed on \gtrsim kiloparsec scales in normal (non-starburst) galaxies. The depletion time of molecular gas can be expressed as $\tau_{\text{H}_2} \equiv \Sigma_{\text{H}_2}/\dot{\Sigma}_* \sim t_{\text{GMC}}/\epsilon_{\text{int}}$, and therefore, if ϵ_{int} scales only with t_{GMC} then τ_{H_2} will be constant, i.e. $\Sigma_{\text{H}_2} \propto \dot{\Sigma}_*$ independent of kiloparsec-scale environment (see Semenov, Kravtsov & Gnedin

2019, for more detail). Indeed, we find that in the region where we apply our analysis, τ_{H_2} is close to constant (see Fig. 3).

4.3 Reconciling the views: molecular cloud evolution driven by the competition between accretion and ejection

We have shown in Section 4.1 that CO-bright molecular clouds are relatively long-lived. By contrast, in Section 4.2, we have shown that their constituent H_2 molecules are very short-lived. In Fig. 13, we show how these two results can be reconciled.

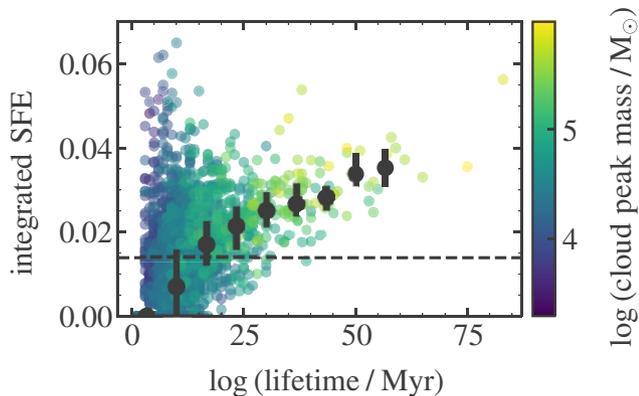


Figure 12. Integrated star formation efficiency over the lifetime of a CO-bright molecular cloud, as a function of its lifetime. The black data points and error bars correspond to the median values and interquartile ranges in each lifetime interval. The horizontal dashed line gives the total stellar mass produced in all clouds, divided by the total mass of all clouds, which is equal to the total SFE within the simulation. The transparent data points correspond to the values for individual molecular clouds, coloured by the peak mass they attain during their lifetimes.

In panel (a), we explicitly show the mass evolution of all molecular clouds in our simulation, as a function of time. To aid visualization, we divide the clouds into five bins of lifetime, indicated by colour. The median values at each time are given by the solid lines and the shaded regions indicate the corresponding interquartile ranges. In agreement with the trend shown in Fig. 7, longer lived molecular clouds achieve higher peak masses. The mass evolution of the molecular clouds in the simulation is remarkably symmetrical. All clouds begin and finish at the same mass, which is determined by the onset of effective H₂ self- and dust-shielding at our mass resolution of 859 M_⊙. The clouds increase in mass for the first half of their lives, and decrease in mass for the second half.

In panel (b) of Fig. 13, we demonstrate that the mass evolution of the molecular clouds can be described by the competing accretion and ejection of gas into and out of a state of high H₂ abundance. The solid lines show the median mass of Lagrangian gas parcels (tracer particles) entering the H₂-dominated state per unit time per molecular cloud; the dashed lines show the total mass exiting this state per unit time. The masses of the molecular clouds (upper panel) switch from increasing to decreasing when the rate of mass ejection crosses above the rate of mass accretion. From mass conservation, the evolution of cloud masses can be described by the simple formula

$$M(t) = \int_0^t dt' [\dot{M}_{\text{accr}}(t') - \dot{M}_{\text{ej}}(t')], \quad (4)$$

where $M(t)$ is the instantaneous mass of the cloud, and \dot{M}_{accr} and \dot{M}_{ej} are the instantaneous rates of H₂ mass accretion and ejection, respectively. The total lifetime of the cloud is therefore given by

$$\int_0^{t_{\text{life}}} dt' \dot{M}_{\text{accr}}(t') = \int_0^{t_{\text{life}}} dt' \dot{M}_{\text{ej}}(t'). \quad (5)$$

We note that this finding is broadly consistent with the analytic models proposed by Goldbaum et al. (2011), Inutsuka et al. (2015), Burkert (2017), and Kobayashi et al. (2017). These authors argue that molecular cloud lifetimes are determined primarily by external accretion rates, and that the galactic distributions of molecular cloud mass and SFR results from a competition between gas accretion driven by the compression of gas at the interfaces between large-

scale converging flows and gas ejection driven by galactic shear and stellar feedback.

Panels (c) and (d) of Fig. 13 demonstrate that the rate of molecular gas ejection from clouds is correlated in time with the instantaneous SFR and, closely related, the momentum injection rate from H II regions. The median rate of star formation per cloud is indicated by the dashed lines in panel (c), and the corresponding radial momentum injected by the H II regions is indicated by the dashed lines in panel (d). The SFR tracks the cloud mass, and therefore the rate of molecular gas ejection by stellar feedback also tracks the cloud mass. It begins at the same value for all clouds, and diverges only after it crosses the mass accretion rate and so begins to destroy the cloud.

We demonstrate the proportionality between the mass ejection rate \dot{M}_{ej} and the rate of momentum injection \dot{p}_r by H II regions more explicitly in Fig. 14. The fact that $\dot{M}_{\text{ej}} \propto \dot{p}_r$ can be understood in terms of the momentum required to eject material from a (spherical) bounding surface at radius r_{shield} around each young star particle, where r_{shield} is the radius at which x_{H_2} drops to 0.5, set by the extent of self- and dust-shielding. In this case,

$$\dot{M}_{\text{ej}} \sim \frac{\dot{p}_r}{v_{\text{esc}}} \sim \frac{1}{r_{\text{shield}}} \sqrt{\frac{3}{8\pi G \rho}} \dot{p}_r, \quad (6)$$

where v_{esc} is the escape speed at r_{shield} , which we write in terms of the mean volume density ρ of the gas inside r_{shield} . The variation in the values of ρ and r_{shield} is not very large, resulting in the proportionality $\dot{M}_{\text{ej}} \propto \dot{p}_r$ shown in Fig. 14.

We therefore find that the rate of mass ejection \dot{M}_{ej} across all molecular clouds in our simulation can be straight forwardly parametrized in terms of the rate of momentum injection \dot{p}_r into these regions by early feedback.

4.4 The clustering of supernovae in massive, long-lived molecular clouds

The concentration of star formation in the most massive and long-lived molecular clouds (Section 4.1), fed by the fast accretion of new molecular gas (Section 4.3), drives the spatial and temporal clustering of supernovae. We demonstrate this in Fig. 15, which shows the two-point correlation function $\xi(\Delta)$ of supernova explosions occurring in more and less massive clouds. The black-dashed line corresponds to the time-averaged median value of the two-point correlation function for all supernova explosions occurring across the simulated galaxy, within all time slices of 1 Myr over the 300-Myr simulation interval. This line indicates that on scales $\Delta \lesssim 80$ pc, the supernovae are *more clustered* ($\xi > 1$) than would be expected for a Poisson (uniform) distribution of objects across the galactic mid-plane, and on scales $\Delta > 80$ pc they are *less clustered* ($\xi < 1$). The degree of clustering of supernovae rises steeply on small scales of $\lesssim 30$ pc, and Fig. 15 demonstrates that this clustering is accounted for almost exclusively by supernovae in clouds with masses $> 10^5 M_{\odot}$ (green line). By contrast, the supernovae in lower mass clouds (purple line) approach a random distribution at small scales. The clustering of supernovae on such small scales, approaching the softening length of 6 pc used in our simulation, has been shown to enhance the momentum injected per supernova to the interstellar medium by a factor of at least 4 (e.g. Gentry et al. 2017). This, in turn, increases the mass-loading of outflows (Fielding, Quataert & Martizzi 2018) as well as the burstiness of star formation across galaxies (Smith et al. 2021).

Fig. 16 helps to visualise this intuitive result, comparing the time-evolving H₂ column density of a low-mass cloud in our simulation

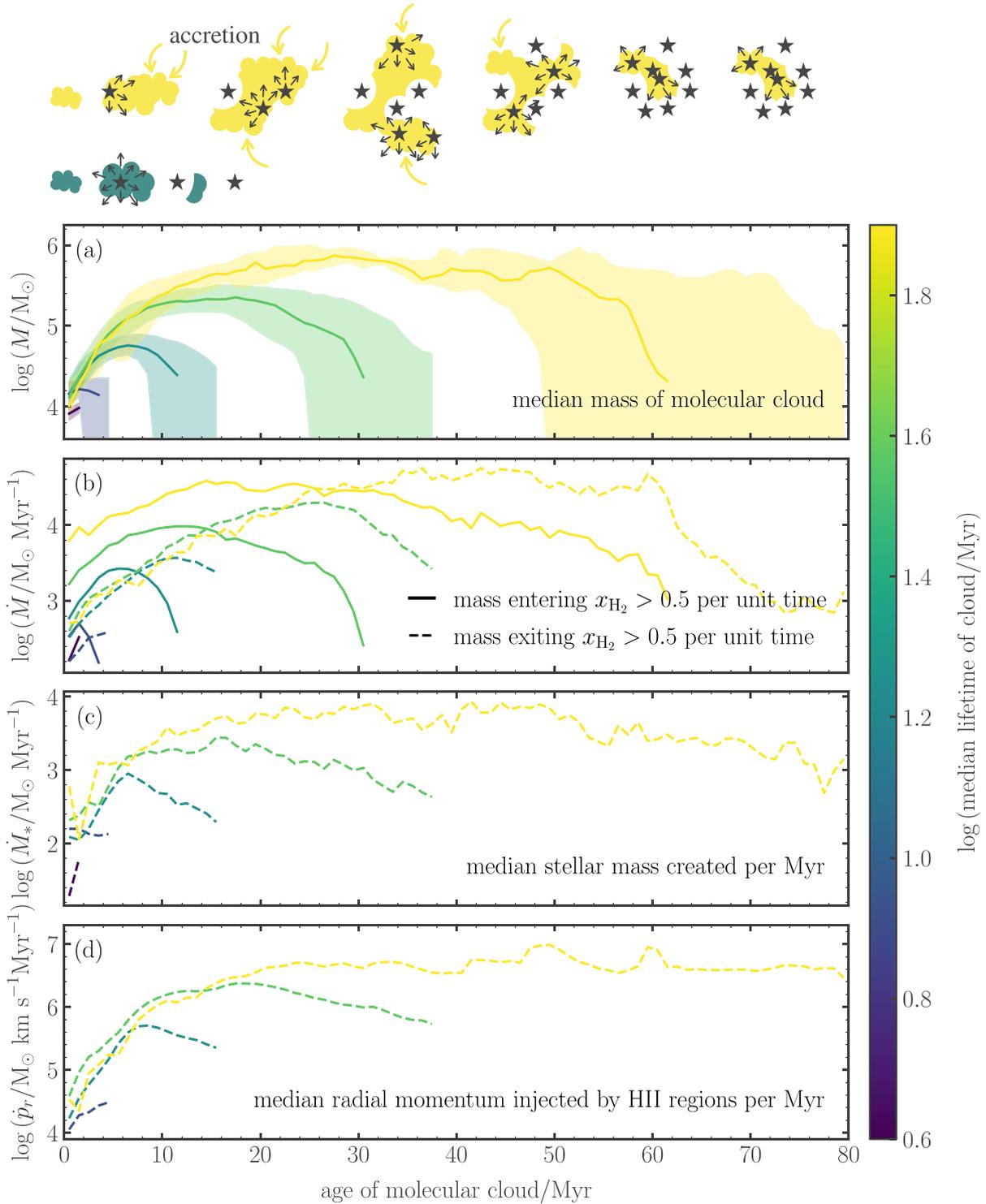


Figure 13. Cloud properties as a function of time since cloud formation, computed as medians over bins of total cloud lifetime, from the shortest-lived clouds (purple) to the longest-lived (yellow). Panel (a) shows instantaneous cloud mass, with the solid lines showing the medians and the shaded bands showing the interquartile range. Panel (b) shows mass flux into (solid) and out of (dashed) the H_2 -dominated phase, $x_{H_2} > 0.5$. Panel (c) shows SFR, and (d) shows rate of radial momentum injection by HII regions; both of these rates are computed using a 5 Myr averaging window to suppress fluctuations due to the stochasticity of star formation. Note that the total mass follows a symmetrical pattern of mass evolution from formation to destruction, and that the mass-loss rate from the H_2 -rich phase, SFR, and momentum injection rate are all strongly correlated. The schematic at the top shows a possible lifecycle for a short-lived (green) and a long-lived (yellow) molecular cloud, including the effects of accretion and feedback-induced gas ejection. The smallest tracer particle population averaged in this figure (lowest cloud lifetime, in purple) is 6.6×10^4 tracer particles.

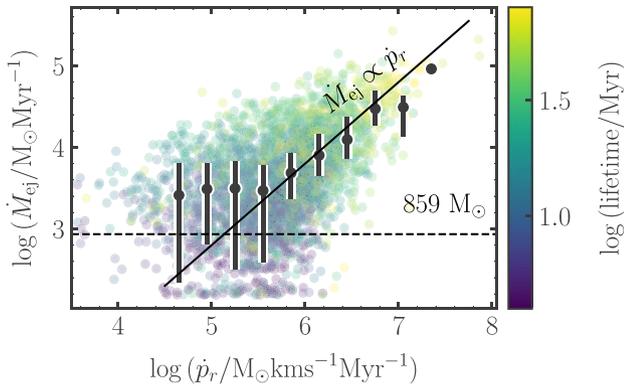


Figure 14. The rate of mass ejection \dot{M}_{ej} as a function of the radial momentum injection rate \dot{p}_r , for all molecular clouds in the simulation that are receiving feedback momentum. The scatter plot shows 1/5th of the total measurements for the cloud population, coloured according to the lifetime of the cloud. The black points and error bars give the median and interquartile range values in 10 different mass bins. To guide the eye, the black solid line corresponds to the case of direct proportionality between mass ejection and momentum injection (not a fit). The black dashed horizontal line shows the resolution limit of the simulation, which corresponds to ejecting one resolution element of mass $\Delta M = 859 M_\odot$ over the time interval $\Delta t = 1$ Myr that we use in our numerical computation of \dot{M}_{ej} .

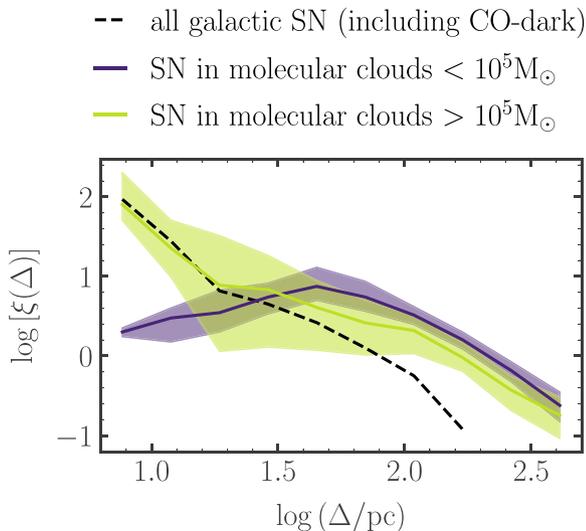


Figure 15. The two-point correlation function $\xi(\Delta)$ for supernova explosions as a function of their separation Δ over time intervals of 1 Myr, averaged over all times (solid lines). The shaded regions give the interquartile ranges over these times. The black dashed line is calculated for all supernovae across the entire simulation, while the green and purple lines represent the supernovae associated with star particles born in long-lived, high-mass and short-lived, low-mass molecular clouds, respectively. Massive and long-lived clouds account for supernova clustering on scales $\lesssim 30$ pc.

($\sim 5 \times 10^4 M_\odot$, upper panels) to that of a high-mass cloud ($\sim 10^6 M_\odot$, lower panels). White-filled stars represent young star particles of age < 5 Myr, which are emitting significant quantities of ionizing radiation, and empty stars denote supernova explosions. The yellow arrows show the velocities of gas tracer particles. In the low-mass cloud, star formation within a radius of ~ 100 pc is halted as the feedback from the young star particle destroys the entire molecular

gas reservoir. By contrast, the high-mass cloud continues to accrete more molecular mass from the lower right-hand corner (inward-pointing yellow arrows), even as an entire cluster of supernovae are occurring in the top left-hand corner. Pre-supernova feedback destroys only small sections of the cloud, while new stars continue to form in other nearby sections. When massive stars eventually explode as supernovae, they are therefore surrounded by other nearby supernova explosions: a cluster of supernovae.

5 DISCUSSION

Here, we discuss both the implications of our results in the context of the existing literature and the caveats associated with our calculations.

5.1 Previous results for molecular cloud lifetimes

Our demonstration that accretion–ejection balance is the key process in molecular cloud evolution helps resolve a number of problems in the literature regarding molecular cloud lifetimes. First, it helps explain why molecular cloud lifetimes computed in simulations of isolated GMCs tend to be too short compared to observations. In such isolated cloud simulations, molecular cloud lifetimes for observationally motivated initial densities are typically 5–10 Myr (e.g. Grudić et al. 2018; He, Ricotti & Geen 2019; Fukushima et al. 2020; Grudić et al. 2021; Kim, Ostriker & Filippova 2021), a factor of 2–5 shorter than the ≈ 20 –30 Myr mean generally inferred from observations of nearby galaxies (e.g. Kawamura et al. 2009; Miura et al. 2012; Kruijssen et al. 2019; Chevance et al. 2020). Our result provides a natural explanation for this discrepancy: Simulations of isolated clouds are necessarily missing the accretion part of the accretion–ejection balance. With no fresh mass supplies, these simulations produce lifetimes closer to the Lagrangian lifetime of individual molecules than to the lifetimes of real molecular clouds. We note that Richings & Schaye (2016) and Koda & Tan (2023) have also recently studied the differences in the lifetimes of star-forming regions inferred via different gas tracers, and by different definitions of these regions.

Our results also help to explain the seeming contradiction between the presence of molecular clouds in the interarm regions of grand design spiral galaxies and a typical observed cloud lifetimes of 20–30 Myr, much less than the interarm transit time. Our simulated dwarf spiral galaxy has an average gas density (1 – 5 cm^{-3}) and velocity dispersion ($\sim 10 \text{ km s}^{-1}$), similar to the values found in the interarm regions of grand design spiral galaxies such as M51. Despite these conditions, and despite the short typical survival time of an H_2 molecule in our simulation (~ 2 –9 Myr, see Figs 9 and 11), we none the less find that massive molecular clouds in our simulation can survive for up to 90 Myr. These long lifetimes are sustained by a high accretion rate of new molecular gas throughout the cloud lifetime: up to $4 \times 10^4 M_\odot \text{ Myr}^{-1}$ for the longest-lived clouds. This accretion rate is an order of magnitude faster than predicted by Koda et al. (2009) and Koda, Scoville & Heyer (2016) for the comparable interarm environment in M51, and therefore provides an explanation for the existence of clouds in this interarm region that is not in contradiction with a much shorter typical molecule survival time (~ 4 Myr) and a much shorter typical cloud survival time (≈ 20 –30 Myr, see Fig. 7). Possible processes driving this fast accretion are converging flows due to supernova-driven bubbles or converging flows due to large-scale galactic-dynamical processes such as galactic shear. Pinpointing and describing this rate of accretion in our simulation will be the topic of a future paper.

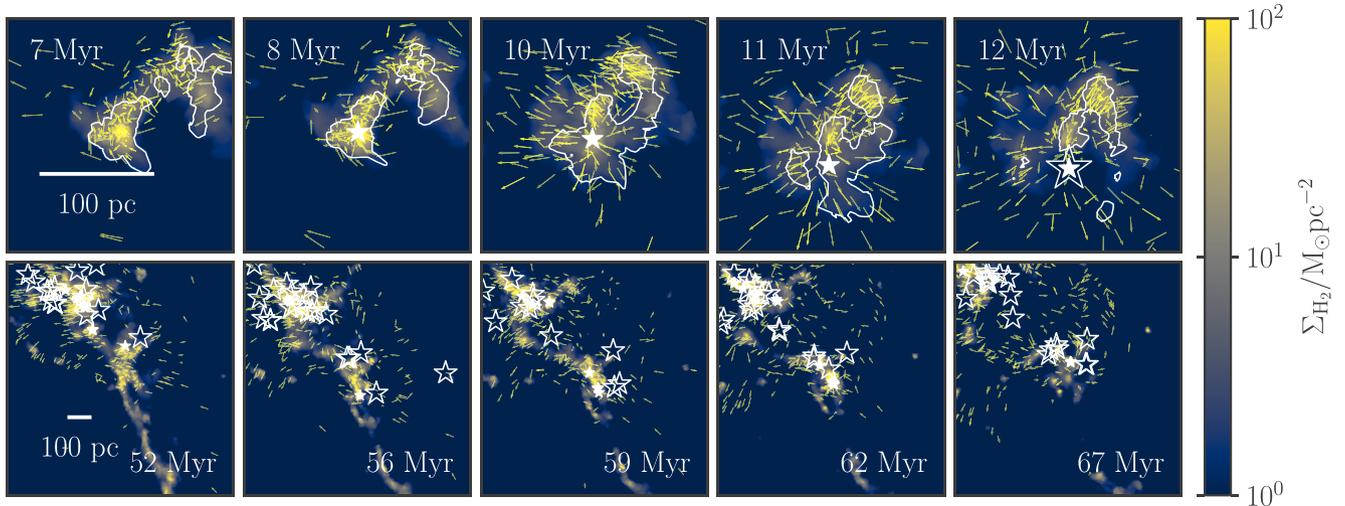


Figure 16. Examples of stellar feedback occurring in a low-mass, short-lived star-forming region ($\sim 5 \times 10^4 M_{\odot}$, upper panels) and in a high-mass, long-lived star-forming region ($\sim 10^6 M_{\odot}$, lower panels). Time runs from left to right across the page, with the region ages given in white. Star particles of age < 5 Myr (emitting pre-supernova feedback) are indicated by filled white stars, while supernova explosions are indicated by open white stars. The velocities of the tracer particles that transit through the clouds are indicated by yellow arrows. The colour bar indicates the total molecular gas surface density, and white contours in the upper panel enclose regions of projected molecular hydrogen abundance $\Sigma_{\text{H}_2}/\Sigma_{\text{g}} \geq 0.3$. High-mass star-forming regions cannot be entirely destroyed by distributed pre-supernova feedback, and so are responsible for the clustering of supernovae on small scales in the simulation.

5.2 Caveats

We conclude this discussion by pointing out two limitations of our calculations, and considering their possible impact.

5.2.1 Unresolved shielding of molecular gas

The resolution of our simulation ($859 M_{\odot}$, with a softening length of 3 pc in the dense molecular gas) presents a potential caveat to the result that the median chemical survival time of H_2 molecules is 4 Myr in a dwarf spiral galaxy. Specifically, the lack of very dense gas clumps within molecular clouds will aid the removal of gas from the cloud and shorten the time-scale for the associated H_2 destruction. Furthermore, Joshi et al. (2019) show that the H_2 abundance derived using our chemical network and shielding prescription is converged only at a resolution of 0.2 pc, due to the existence of sub-pc dense clumps of molecular gas.

In such clumps, both the ratio of the H_2 formation and destruction rates and the amount of self-shielding would be enhanced compared to those, we measure at our ≈ 3 pc resolution, and as a result the value of 4 Myr that we have derived for this lifetime may be an underestimate. This would also mean that the rate \dot{M}_{ej} of gas ejection from molecular clouds would be lowered relative to the value we have calculated. The rate \dot{M}_{accr} of gas accretion would be unaffected. However, there is circumstantial evidence that this cannot be a major effect. We have shown in Fig. 3 that the average H I and H_2 gas densities and depletion times in our simulated galaxy are in good agreement with observations, as are the masses and sizes of its giant molecular clouds (Fig. 5). If the H_2 survival time were significantly longer, agreement with observations would worsen. This suggests that the rate limiting step in H_2 destruction is not the time-scale for chemical reactions, which we do not resolve, but the time-scale for molecular material to be dynamically ejected from well-shielded clouds, which we do. Thus, our failure to capture the details of the chemistry does minimal harm. Given this discussion, however, we do caution that the median H_2 survival time of 4 Myr depends on

a number of factors, including the resolution of our simulation, the properties of the large-scale galactic environment, and the metallicity of the simulated gas reservoir (see the next section). We have shown that this survival time is short, but great significance should not be attached to its exact value. The exact survival time of H_2 molecules may be longer than we have found, due to the longer survival times of unresolved gas clumps, but our main qualitative conclusions remain unaffected.

5.2.2 Galactic-scale metallicity variations

We have assumed a globally invariant, Solar value for the gas-to-dust ratio in our simulation. However, the observed metallicity of NGC300 in fact varies between 0.8 and 0.2 times the Solar value, decreasing monotonically with galactocentric radius (Bresolin et al. 2009). According to our competition model for the evolution of molecular clouds, lowering the gas-phase metallicity in the outer regions of the galactic disc may have several different (competing) effects. Reducing the metallicity will reduce the shielding of H_2 from the ISRF, and so reduce the distance from young stars at which H_2 is ionized as it is ejected by early stellar feedback. This will cause an increase in the rate \dot{M}_{ej} of H_2 ejection from molecular clouds. However, lowered metallicity will also reduce the degree of cooling and star formation in the dense gas, causing a decrease in \dot{M}_{ej} . The effect of lowered metallicity on the mass accretion rate \dot{M}_{accr} is more straight-forward: (1) the dust-catalyzed formation of H_2 from atomic hydrogen will be slower, and (2) gas parcels will need to reach higher densities before becoming shielded from the ISRF. Both of these effects will act to decrease the rate of H_2 accretion on to molecular clouds. As such, it is difficult to predict exactly how the inclusion of live metal-enrichment will affect our results, except to note that the environmentally dependent metallicity will likely be a variable that plays a significant role in setting \dot{M}_{ej} . Further studies that vary the metallicity and control for galaxy morphology are necessary to determine the functional form of this dependence.

6 CONCLUSIONS

In Section 1, we posed two questions: (1) What is the chemical lifetime of H₂ in the interstellar medium of galaxies and how does it relate to the molecular cloud lifetime? And (2) how can we reconcile the long observed lifetimes of giant molecular clouds (e.g. Koda et al. 2009) and the short Lagrangian residence times of gas in the star-forming state (e.g. Semenov, Kravtsov & Gnedin 2017)? In this work, we have used passive gas tracer particles in an AREPO simulation of a dwarf spiral galaxy to arrive at the following answers:

(i) The typical survival time of H₂ molecules is 4 Myr, with an interquartile range from 2 to 9 Myr in our simulation, subject to the accuracy of our chemical model. This H₂ destruction time-scale is independent of the lifetime of the host molecular cloud, which has a much longer H₂ mass-weighted median value of ~ 25 Myr. The rapid destruction of H₂ molecules is driven by early stellar feedback from H II regions.

(ii) The reason that molecular clouds are able to survive much longer than their constituent molecules is that the molecular cloud lifetime is determined by the competition between the feedback-driven ejection of molecular gas, and the accretion of new molecular gas from the large-scale galactic environment. This competition means that long-lived and massive molecular clouds can be sustained given a sufficiently fast accretion rate. In our simulation, a value of $\dot{M}_{\text{accr}} \sim 4 \times 10^4 M_{\odot} \text{ Myr}^{-1}$ can feed a molecular cloud for 90 Myr, and up to a peak mass of $10^6 M_{\odot}$.

Our answers have important implications for the dynamics of gas on both cloud and galactic scales. In particular, we find that

(i) The time spent by gas in the H₂-dominated state is much shorter than the time-spent in the H₂-poor state, in qualitative agreement with the results of Semenov, Kravtsov & Gnedin (2017). A given Lagrangian parcel of gas spends most of its life transiting between molecular clouds, with only brief passages through clouds.

(ii) A partial exception to this is that gas can sometimes become ‘trapped’ around higher mass molecular clouds, cycling back to a dense, H₂-dominated state faster and more frequently than occurs for gas parcels passing through low-mass clouds. This ‘trapping’ raises the integrated star formation efficiency in massive, long-lived molecular clouds up to four times higher than that in low-mass, short-lived clouds.

(iii) Despite their small number, massive and long-lived molecular clouds account for a not insignificant amount of the total star formation budget of a galaxy. Moreover, they drive the clustering of supernova feedback on sub-cloud scales, which is in turn a key driver of galactic outflows.

In an upcoming paper, we will investigate the galactic-scale physics driving the fast accretion of new gas on to massive, long-lived molecular clouds, and so driving their higher star formation efficiencies and levels of supernova clustering.

ACKNOWLEDGEMENTS

We thank the anonymous referee for an insightful and attentive report that improved the communication of the results presented in this manuscript. We are very grateful to Shy Genel, Greg Bryan, Blakesley Burkhart, Enrique Vázquez-Semadeni, Charles Lada, Shyam Menon, and Munan Gong for helpful discussions. We thank Volker Springel for providing us access to Arepo. SMRJ is supported by Harvard University through the Institute for Theory and Computation Fellowship. Support for VAS was provided by NASA

through the NASA Hubble Fellowship grant HST-HF-2-51445.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555, and by Harvard University through the Institute for Theory and Computation Fellowship. MRK acknowledges support from the Australian Research Council through Laurate Fellowship FL220100020. The work was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI; award jh2), which is supported by the Australian Government.

DATA AVAILABILITY

All data underlying this work can be found in the article.

REFERENCES

- Bertoldi F., McKee C. F., 1992, *ApJ*, 395, 140
 Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, 136, 2846
 Blitz L., Fukui Y., Kawamura A., Leroy A., Mizuno N., Rosolowsky E., 2007, in Reipurth B., Jewitt D., Keil K., eds, *Protostars and Planets V*. Univ. Arizona Press, Tucson, AZ, p. 81
 Bolatto A. D., Wolfire M., Leroy A. K., 2013, *ARA&A*, 51, 207
 Braun H., Schmidt W., 2015, *MNRAS*, 454, 1545
 Bresolin F., Gieren W., Kudritzki R.-P., Pietrzyński G., Urbaneja M. A., Carraro G., 2009, *ApJ*, 700, 309
 Burkert A., 2017, *Mem. Soc. Astron. Italiana*, 88, 533
 Chabrier G., 2003, *Publ. Astron. Soc. Pac.*, 115, 763
 Chevance M. et al., 2020, *MNRAS*, 493, 2872
 Clark P. C., Glover S. C. O., Klessen R. S., 2012, *MNRAS*, 420, 745
 Colombo D. et al., 2019, *MNRAS*, 483, 4291
 Corbelli E. et al., 2017, *A&A*, 601, A146
 da Silva R. L., Fumagalli M., Krumholz M., 2012, *ApJ*, 745, 145
 da Silva R. L., Fumagalli M., Krumholz M. R., 2014, *MNRAS*, 444, 3275
 Daddi E. et al., 2010, *ApJ*, 714, L118
 Dobbs C. L., Pringle J. E., 2013, *MNRAS*, 432, 653
 Engargiola G., Plambeck R. L., Rosolowsky E., Blitz L., 2003, *ApJS*, 149, 343
 Faesi C. M., Lada C. J., Forbrich J., 2018, *ApJ*, 857, 19
 Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994a, *A&AS*, 104, 365
 Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994b, *A&AS*, 105, 29
 Federrath C., Banerjee R., Clark P. C., Klessen R. S., 2010, *ApJ*, 713, 269
 Fielding D., Quataert E., Martizzi D., 2018, *MNRAS*, 481, 3325
 Fujimoto Y., Chevance M., Haydon D. T., Krumholz M. R., Kruijssen J. M. D., 2019, *MNRAS*, 487, 1717
 Fukushima H., Yajima H., Sugimura K., Hosokawa T., Omukai K., Matsumoto T., 2020, *MNRAS*, 497, 3830
 Genel S., Vogelsberger M., Nelson D., Sijacki D., Springel V., Hernquist L., 2013, *MNRAS*, 435, 1426
 Gensior J., Kruijssen J. M. D., Keller B. W., 2020, *MNRAS*, 495, 199
 Gentry E. S., Krumholz M. R., Dekel A., Madau P., 2017, *MNRAS*, 465, 2471
 Glover S. C. O., Mac Low M.-M., 2007a, *ApJS*, 169, 239
 Glover S. C. O., Mac Low M.-M., 2007b, *ApJ*, 659, 1317
 Glover S. C. O., Federrath C., Mac Low M. M., Klessen R. S., 2010, *MNRAS*, 404, 2
 Goldbaum N. J., Krumholz M. R., Matzner C. D., McKee C. F., 2011, *ApJ*, 738, 101
 Gong M., Ostriker E. C., Wolfire M. G., 2017, *ApJ*, 843, 38
 Gratier P. et al., 2012, *A&A*, 542, A108
 Greif T. H., Springel V., White S. D. M., Glover S. C. O., Clark P. C., Smith R. J., Klessen R. S., Bromm V., 2011, *ApJ*, 737, 75
 Grisdale K., Agertz O., Renaud F., Romeo A. B., Devriendt J., Slyz A., 2019, *MNRAS*, 486, 5482

Grudić M. Y., Hopkins P. F., Faucher-Giguère C.-A., Quataert E., Murray N., Kereš D., 2018, *MNRAS*, 475, 3511

Grudić M. Y., Kruijssen J. M. D., Faucher-Giguère C.-A., Hopkins P. F., Ma X., Quataert E., Boylan-Kolchin M., 2021, *MNRAS*, 506, 3239

He C.-C., Ricotti M., Geen S., 2019, *MNRAS*, 489, 1880

Hollyhead K., Bastian N., Adamo A., Silva-Villa E., Dale J., Ryon J. E., Gazak Z., 2015, *MNRAS*, 449, 1106

Hopkins P. F., Narayanan D., Murray N., 2013, *MNRAS*, 432, 2647

Hopkins P. F. et al., 2018, *MNRAS*, 480, 800

Hughes A. et al., 2013, *ApJ*, 779, 46

Inutsuka S.-i., Inoue T., Iwasaki K., Hosokawa T., 2015, *A&A*, 580, A49

Jeffreson S. M. R., Kruijssen J. M. D., Keller B. W., Chevance M., Glover S. C. O., 2020, *MNRAS*, 498, 385

Jeffreson S. M. R., Keller B. W., Winter A. J., Chevance M., Kruijssen J. M. D., Krumholz M. R., Fujimoto Y., 2021a, *MNRAS*, 505, 1678

Jeffreson S. M. R., Krumholz M. R., Fujimoto Y., Armillotta L., Keller B. W., Chevance M., Kruijssen J. M. D., 2021b, *MNRAS*, 505, 3470

Joshi P. R., Walch S., Seifried D., Glover S. C. O., Clarke S. D., Weis M., 2019, *MNRAS*, 484, 1735

Kawamura A. et al., 2009, *ApJS*, 184, 1

Kennicutt R. C., Jr, 1998, *ARA&A*, 36, 189

Kim J.-G., Ostriker E. C., Filippova N., 2021, *ApJ*, 911, 128

Kim J. et al., 2022, *MNRAS*, 516, 3006

Kimm T., Cen R., 2014, *ApJ*, 788, 121

Kobayashi M. I. N., Inutsuka S.-i., Kobayashi H., Hasegawa K., 2017, *ApJ*, 836, 175

Koda J., Tan J. C., 2023, preprint ([arXiv:2308.11717](https://arxiv.org/abs/2308.11717))

Koda J. et al., 2009, *ApJ*, 700, L132

Koda J., Scoville N., Heyer M., 2016, *ApJ*, 823, 76

Kretschmer M., Teyssier R., 2020, *MNRAS*, 492, 1385

Kruijssen J. M. D. et al., 2019, *Nature*, 569, 519

Krumholz M. R., 2013, Astrophysics Source Code Library, record ascl:1304.007

Krumholz M. R., 2014, *MNRAS*, 437, 1662

Krumholz M. R., Matzner C. D., 2009, *ApJ*, 703, 1352

Krumholz M. R., Fumagalli M., da Silva R. L., Rendahl T., Parra J., 2015, *MNRAS*, 452, 1447

Lee E. J., Miville-Deschênes M.-A., Murray N. W., 2016, *ApJ*, 833, 229

Leitherer C. et al., 1999, *ApJS*, 123, 3

Leroy A. K. et al., 2013, *AJ*, 146, 19

MacLaren I., Richardson K. M., Wolfendale A. W., 1988, *ApJ*, 333, 821

Mathis J. S., Mezger P. G., Panagia N., 1983, *A&A*, 500, 259

Matzner C. D., 2002, *ApJ*, 566, 302

Miura R. E. et al., 2012, *ApJ*, 761, 37

Miville-Deschênes M.-A., Murray N., Lee E. J., 2017, *ApJ*, 834, 57

Mooney T. J., Solomon P. M., 1988, *ApJ*, 334, L51

Murray N., 2011, *ApJ*, 729, 133

Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493

Nelson A. F., 2006, *MNRAS*, 373, 1039

Nelson R. P., Langer W. D., 1997, *ApJ*, 482, 796

Onodera S. et al., 2010, *ApJ*, 722, L127

Padoan P., Haugbølle T., Nordlund Å., Frimann S., 2017, *ApJ*, 840, 48

Pineda J. E., Rosolowsky E. W., Goodman A. A., 2009, *ApJ*, 699, L134

Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, *MNRAS*, 338, 14

Richings A. J., Schaye J., 2016, *MNRAS*, 460, 2297

Robertson B. E., Kravtsov A. V., 2008, *ApJ*, 680, 1083

Rosolowsky E. et al., 2021, *MNRAS*, 502, 1218

Safranek-Shrader C., Krumholz M. R., Kim C.-G., Ostriker E. C., Klein R. I., Li S., McKee C. F., Stone J. M., 2017, *MNRAS*, 465, 885

Schruba A., Leroy A. K., Walter F., Sandstrom K., Rosolowsky E., 2010, *ApJ*, 722, 1699

Schruba A. et al., 2011, *AJ*, 142, 37

Scoville N. Z., Hersch K., 1979, *ApJ*, 229, 578

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

Semenov V. A., Kravtsov A. V., Gnedin N. Y., 2021, *ApJ*, 918, 13

Shin E.-j., Tacchella S., Kim J.-h., Iyer K. G., Semenov V. A., 2022, *ApJ*, 947, 18

Smith R. J., Glover S. C. O., Clark P. C., Klessen R. S., Springel V., 2014, *MNRAS*, 441, 1628

Smith M. C., Bryan G. L., Somerville R. S., Hu C.-Y., Teyssier R., Burkhardt B., Hernquist L., 2021, *MNRAS*, 506, 3882

Sokal K. R., Johnson K. E., Indebetouw R., Massey P., 2016, *ApJ*, 826, 194

Solomon P. M., Vanden Bout P. A., 2005, *ARA&A*, 43, 677

Springel V., 2010, *MNRAS*, 401, 791

Sun J. et al., 2018, *ApJ*, 860, 172

Teyssier R., 2015, *ARA&A*, 53, 325

Tress R. G., Smith R. J., Sormani M. C., Glover S. C. O., Klessen R. S., Mac Low M.-M., Clark P. C., 2020, *MNRAS*, 492, 2973

Truelove J. K., Klein R. I., McKee C. F., Holliman J. H. II, Howell L. H., Greenough J. A., 1997, *ApJ*, 489, L179

van der Tak F. F. S., van Dishoeck E. F., 2000, *A&A*, 358, L79

Vázquez G. A., Leitherer C., 2005, *ApJ*, 621, 695

Westmeier T., Braun R., Koribalski B. S., 2011, *MNRAS*, 410, 2217

Wong T., Blitz L., 2002, *ApJ*, 569, 157

Wyder T. K. et al., 2009, *ApJ*, 696, 1834

APPENDIX A: CONVERGENCE OF THE INITIAL TRACER MASS DISTRIBUTION

In Fig. A1, we demonstrate that the initial, non-uniform effective mass distribution of tracer particles in our simulation converges to a uniform distribution on a time-scale of <100 Myr. As noted in Section 2.3, we assign one tracer particle per gas cell in our initial condition, and the initial spread of gas cell masses is over an order of magnitude. This means that the initial spread of effective tracer particle masses is over an order of magnitude.

In order to demonstrate that this initial non-uniform distribution rapidly approaches one that samples uniformly in mass, in Fig. A1, we plot the number of tracer particles as a function of gas surface density computed in 2D columns through the galactic mid-plane, at a resolution of 6 pc. The black solid line has slope unity, and shows the

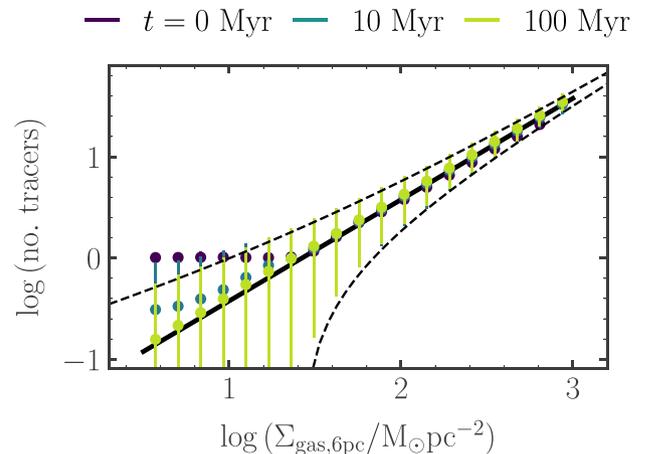


Figure A1. Mean and standard deviation of the tracer number as a function of gas surface density in 2D columns of area $(6 \text{ pc})^2$ through the entire mid-plane of the simulated galaxy disc, at simulation times of 0, 10, and 100 Myr. The pixel size corresponds to the native resolution of the simulation, at which molecular clouds are identified and analysed in two dimensions. The thick black line corresponds to one tracer particle per $450 M_{\odot}$, while the dashed black lines show a range of $\pm\sqrt{\text{no. tracers}}$ around this, the error expected from Poisson sampling for $\text{no. tracers} \gg 1$.

expected relationship for a uniform distribution of effective tracer particle masses, while the dashed lines denote the Poisson error expected due to the Monte Carlo exchange of a small number of tracer particles between gas cells. The plot shows that, after a period of 100 Myr (green points and error bars), both the number of tracer particles per pixel, and the error in this number, conform to the expectation for a uniform distribution of effective tracer particle masses.

APPENDIX B: CHEMICAL POST-PROCESSING

As noted in Section 3, our CO-luminous star-forming regions are identified using two-dimensional maps of the CO-bright molecular gas column density, $\Sigma_{H_2,CO}$. To calculate this column density, we post-process the simulation output using the DESPOTIC model for astrochemistry and radiative transfer (Krumholz 2013, 2014), following the method outlined in Fujimoto et al. (2019) except as noted below; we summarize the procedure here, and refer readers to Fujimoto et al. for further details. The self- and dust-shielding of CO molecules from the ambient UV radiation field, and the effects of non-local thermodynamic equilibrium excitation and varying CO line optical depth, cannot be accurately computed during runtime at the mass resolution of our simulation. Within DESPOTIC, the escape probability formalism is applied to compute the CO line emission from each gas cell according to its hydrogen atom number density n_H , column density N_H and virial parameter α_{vir} , assuming that the cells are approximately spherical. In practice, the line luminosity varies smoothly with the variables n_H , N_H , and α_{vir} . We therefore interpolate over a grid of pre-calculated models at regularly spaced logarithmic intervals in these variables to reduce computational cost. The hydrogen column density is estimated via the local approximation of Safronek-Shrader et al. (2017) as $N_H = \lambda_J n_H$, where $\lambda_J = (\pi c_s^2 / G \rho)^{1/2}$ is the Jeans length, with an upper limit of $T = 40$ K on the gas cell temperature. The virial parameter is calculated from the turbulent velocity dispersion of each gas cell according to MacLaren, Richardson & Wolfendale (1988), and Bertoldi & McKee (1992). The line emission is self-consistently coupled to the chemical and thermal evolution of the gas, including carbon and oxygen chemistry (Gong, Ostriker & Wolfire 2017), gas heating by cosmic rays and the grain photoelectric effect, line cooling due to C^+ , C, O, and CO and thermal exchange between dust and gas. We match the ISRF strength and cosmic-ray ionization rate to the values used in our live chemistry.

Having calculated values of the CO line luminosity for each simulated gas cell, we compute the CO-bright molecular hydrogen surface density as

$$\Sigma_{H_2,CO} [M_\odot pc^{-2}] = \frac{2.3 \times 10^{-29} M_\odot (erg s^{-1})^{-1}}{m_H [M_\odot]} \times \int_{-\infty}^{\infty} dz' \rho_g(z') L_{CO} [erg s^{-1} H atom^{-1}], \quad (B1)$$

where $\rho_g(z)$ is the total gas volume density in $M_\odot pc^{-3}$ at a distance z (in pc) from the galactic mid-plane. The factor of $2.3 \times 10^{-29} M_\odot (erg s^{-1})^{-1}$ combines the mass-to-luminosity conversion factor $\alpha_{CO} = 4.3 M_\odot pc^{-2} (K kms^{-1})^{-1}$ of Bolatto, Wolfire & Leroy (2013) with the line-luminosity conversion factor $5.31 \times 10^{-30} (K km s^{-1} pc^2) / (erg s^{-1})$ for the CO $J = 1 \rightarrow 0$ transition at redshift $z = 0$ (Solomon & Vanden Bout 2005). The physical meaning of $\Sigma_{H_2,CO}$ is simply that it is the H_2 surface density that one would infer from an observation in the CO $J = 1 \rightarrow 0$ line assuming a fixed conversion factor $\alpha_{CO} = 4.3 M_\odot pc^{-2} (K kms^{-1})^{-1}$.

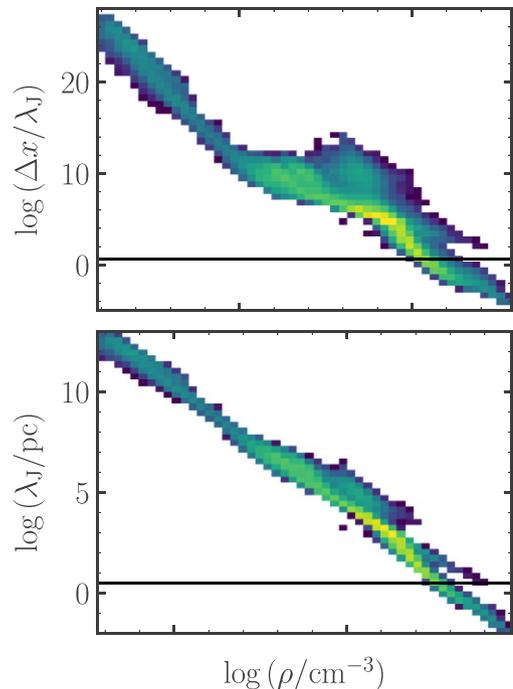


Figure C1. Upper panel: gas cell size Δx divided by the Jeans length λ_J for all gas cells in our simulation, as a function of gas density ρ , and weighted by gas mass. The black horizontal line gives the resolution criterion of Truelove et al. (1997). Lower panel: the Jeans length for all gas cells in our simulation, as a function of gas density ρ . The black horizontal line gives the minimum softening length of 3 pc in our simulations.

APPENDIX C: DISCUSSION OF ARTIFICIAL FRAGMENTATION

We noted in Section 2.4 that we use the adaptive gravitational softening prescription in AREPO to avoid the majority of artificial fragmentation, as discussed in Nelson (2006) and Jeffreson et al. (2020). Here, we include a fuller discussion of this problem, in the context of our simulated galaxy.

Broadly speaking, there are two schools of thought on the prevention of artificial, numerically induced fragmentation in gas for which the Jeans length is unresolved (e.g. Truelove et al. 1997; Greif et al. 2011). The first is to remove all gas with an unresolved Jeans length, by introducing a pressure floor (e.g. Robertson & Kravtsov 2008). This has the advantage of preventing artificial fragmentation, but the disadvantage of suppressing physical, gravitational fragmentation in high-density gas, inflating the Jeans length to unphysically high values (see e.g. Teyssier 2015; Hopkins et al. 2018).

The second approach, which we have adopted, is to allow high-density, low-pressure gas, but to take advantage of adaptive gravitational softening and a gravitation-dependent star formation efficiency to avoid the majority of artificial fragmentation in this gas (e.g. Federrath et al. 2010; Hopkins et al. 2018; Tress et al. 2020). The top panel of Fig. C1 presents a 2D histogram of the Truelove et al. (1997) Jeans number in our simulation, as a function of the gas density. This demonstrates that some dense molecular gas is in fact below the resolution limit defined in that work, given by the black horizontal line. However, the lower panel of the figure demonstrates that the Jeans length for all of this dense gas is below the softening length of the simulation, which is 3 pc (black horizontal line). This gas is additionally star forming ($\rho > 100 cm^{-3}$, see Section 2.4).

Perturbations on these scales are therefore unlikely to lead to artificially induced gravitational collapse and fragmentation, though to our knowledge, this has never been thoroughly investigated in the case of adaptive gravitational softening.

Finally, we note that no work has yet performed a thorough investigation of artificial fragmentation in a simulation comparable to ours. The paper of Truelove et al. (1997) concerns an adaptive-mesh code (rather than a moving-mesh code like AREPO), under isothermal conditions with no turbulence, no rotating disc environment, and no gravitational softening, let alone adaptive softening. The work of Nelson (2006) partly addresses two of these problems by introducing a disc environment and adaptive gravitational softening, but looks only at the case of an SPH code, and focuses on the case of protoplanetary discs, which do not have substantial supersonic

turbulence. Finally, Greif et al. (2011) investigate the problem in AREPO, but again in an isothermal box with no turbulence or rotating disc environment. Gravitational softening is used in the latter work, but it is not adaptive.

It would be of substantial benefit to the community to investigate the onset of artificial fragmentation in turbulent, rotating galactic environments with adaptive gravitational softening, as is the subject of many current high-resolution simulations of the interstellar medium.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.