

# Star formation and molecular hydrogen in dwarf galaxies: a non-equilibrium view

Chia-Yu Hu,<sup>1★</sup> Thorsten Naab,<sup>1</sup> Stefanie Walch,<sup>2</sup> Simon C. O. Glover<sup>3</sup>  
and Paul C. Clark<sup>4</sup>

<sup>1</sup>Max-Planck-Institut für Astrophysik, Karl-Schwarzschild Strasse 1, D-85740 Garching, Germany

<sup>2</sup>Physikalisches Institut der Universität zu Köln, Zùlpicher Strasse 77, D-50937 Köln, Germany

<sup>3</sup>Zentrum für Astronomie der Universität Heidelberg, Institut für Theoretische Astrophysik, Albert-Ueberle-Str 2, D-69120 Heidelberg, Germany

<sup>4</sup>School of Physics and Astronomy, Cardiff University, 5 The Parade, Cardiff CF24 3AA, UK

Accepted 2016 March 3. Received 2016 February 29; in original form 2015 October 19

## ABSTRACT

We study the connection of star formation to atomic (H I) and molecular hydrogen (H<sub>2</sub>) in isolated, low-metallicity dwarf galaxies with high-resolution ( $m_{\text{gas}} = 4 M_{\odot}$ ,  $N_{\text{ngb}} = 100$ ) smoothed particle hydrodynamics simulations. The model includes self-gravity, non-equilibrium cooling, shielding from a uniform and constant interstellar radiation field, the chemistry of H<sub>2</sub> formation, H<sub>2</sub>-independent star formation, supernova feedback and metal enrichment. We find that the H<sub>2</sub> mass fraction is sensitive to the adopted dust-to-gas ratio and the strength of the interstellar radiation field, while the star formation rate is not. Star formation is regulated by stellar feedback, keeping the gas out of thermal equilibrium for densities  $n < 1 \text{ cm}^{-3}$ . Because of the long chemical time-scales, the H<sub>2</sub> mass remains out of chemical equilibrium throughout the simulation. Star formation is well correlated with cold ( $T \leq 100 \text{ K}$ ) gas, but this dense and cold gas – the reservoir for star formation – is dominated by H I, not H<sub>2</sub>. In addition, a significant fraction of H<sub>2</sub> resides in a diffuse, warm phase, which is not star-forming. The interstellar medium is dominated by warm gas ( $100 \text{ K} < T \leq 3 \times 10^4 \text{ K}$ ) both in mass and in volume. The scaleheight of the gaseous disc increases with radius while the cold gas is always confined to a thin layer in the mid-plane. The cold gas fraction is regulated by feedback at small radii and by the assumed radiation field at large radii. The decreasing cold gas fractions result in a rapid increase in depletion time (up to 100 Gyr) for total gas surface densities  $\Sigma_{\text{H I}+\text{H}_2} \lesssim 10 M_{\odot} \text{ pc}^{-2}$ , in agreement with observations of dwarf galaxies in the Kennicutt–Schmidt plane.

**Key words:** methods: numerical – galaxies: evolution – galaxies: ISM.

## 1 INTRODUCTION

Dwarf galaxies are thought to be the building blocks of larger galaxies in the hierarchical picture of galaxy formation. While they contribute little to the total mass budget of the galaxy population, they are the most numerous type of galaxy in the local Universe. Dwarf galaxies tend to contain fewer heavy elements compared to more massive ones (Tremonti et al. 2004; Hunter et al. 2012). As such, they are ideal laboratories for studying physical processes of the interstellar medium (ISM) under chemically simple conditions, which can be quite different from those found in normal spiral galaxies such as the Milky Way.

On kpc-scales, spatially resolved observations of nearby star-forming spiral galaxies have demonstrated that the surface density of the star formation rate (SFR) has a clear correlation with the molecular hydrogen (H<sub>2</sub>) surface density and little correlation with the atomic hydrogen (H I) surface density (Bigiel et al. 2008, 2011). In light of the observational evidence, H<sub>2</sub>-dependent sub-resolution recipes for star formation have been widely adopted in hydrodynamical simulations as well as semi-analytic models, where the H<sub>2</sub> abundances are calculated either by directly incorporating non-equilibrium chemistry models (e.g. Gnedin, Tassis & Kravtsov 2009; Christensen et al. 2012) or by analytical ansatz assuming chemical equilibrium (e.g. Fu et al. 2010; Kuhlen et al. 2012; Hopkins et al. 2014; Thompson et al. 2014; Somerville, Popping & Trager 2015). The implicit assumption is that star formation only takes place in H<sub>2</sub>-dominated clouds, and the H I-to-H<sub>2</sub> transition has been thought to be responsible for the so-called star formation

\* E-mail: [huchiayu@mpa-garching.mpg.de](mailto:huchiayu@mpa-garching.mpg.de)

threshold, the surface density of gas below which star formation becomes extremely inefficient.

However, recent theoretical studies have cast doubt on the causal connection between  $H_2$  and star formation (Krumholz, Leroy & McKee 2011; Glover & Clark 2012a), based on the insensitivity of the gas thermal properties to the  $H_2$  abundances. Radiative cooling from  $H_2$  is negligible when the gas temperature drops below a few hundred Kelvin. The formation of carbon monoxide (CO), which is indeed an efficient coolant at very low temperatures, does require the presence of  $H_2$ . Nevertheless, singly ionized carbon ( $C^+$ ) provides cooling that is almost equally efficient as CO at all but the lowest temperatures through fine structure line emission. It is therefore possible for gas devoid of both  $H_2$  and CO to cool down to low enough temperatures and form stars by gravitational collapse. In this picture, the correlation between  $H_2$  and star formation originates from the fact that both  $H_2$  formation and star formation take place in regions well shielded from the interstellar radiation field (ISRF) instead of the former being a necessary condition for the latter. Such a correlation is expected to break down in low-metallicity environments where the  $H_2$  formation time-scales are much longer. Indeed, Krumholz (2012) predicts that star formation would proceed in  $H\text{I}$ -dominated clouds if the gas metallicity is below a few per cent of solar metallicity ( $Z_\odot$ ) based on time-scale arguments. Glover & Clark (2012c) also examine this issue, using simulations of isolated clouds, and show that the SFR of the clouds is insensitive to their molecular content and that for metallicities  $Z \sim 0.1 Z_\odot$  and below, the star-forming clouds are dominated by atomic gas.

Observationally, the detection of  $H_2$  in star-forming dwarf galaxies has proven to be very challenging. CO emission, as the standard approach to derive the  $H_2$  abundance, tends to be very faint in these galaxies. Schrubba et al. (2012) observed 16 nearby star-forming dwarf galaxies and only detected CO successfully in five galaxies with oxygen abundance  $12 + \log_{10}(\text{O}/\text{H}) \gtrsim 8.0$  with very low CO luminosities per SFR compared to those found in massive spiral galaxies. CO was not detected in the other 11 galaxies with  $12 + \log_{10}(\text{O}/\text{H}) \lesssim 8.0$  even if stacking techniques were used. The interpretation was that the CO-to- $H_2$  conversion factor is significantly higher in low-metallicity environment, under the assumption that there should be much more  $H_2$  given their SFR (i.e. assuming constant  $H_2$  depletion time  $\approx 2$  Gyr). The dependence of the CO-to- $H_2$  conversion factor on metallicity is already observed in Local Group galaxies using dust modelling to estimate the  $H_2$  mass (Leroy et al. 2011) and has considerable theoretical support (Wolfire, Hollenbach & McKee 2010; Glover & Mac Low 2011; Shetty et al. 2011a,b; Bolatto, Wolfire & Leroy 2013). However, an alternative interpretation is that  $H_2$  is also rare and star formation in these systems is taking place in regions dominated by atomic hydrogen (see Michałowski et al. 2015, who adopted such an interpretation for their observed galaxies).

In this paper, we conduct numerical simulations to study the ISM properties in an isolated star-forming dwarf galaxy. Our model includes gravity, hydrodynamics, non-equilibrium cooling and chemistry, shielding from the ISRF, an  $H_2$ -independent star formation recipe, stellar feedback and metal enrichment in a self-consistent manner. We investigate the relationship between  $H_2$  and star formation and explore the effects of varying the strength of ISRF and the dust-to-gas mass ratio (DGR). In Section 2, we describe the details of our numerical method. In Section 3, we show the ISM properties when it is in thermal and chemical equilibrium. In Section 4, we present the results of our numerical simulations. In Section 5, we discuss the implications of our results and the potential caveats. In Section 6, we summarize our work.

## 2 NUMERICAL METHOD

### 2.1 Gravity and hydrodynamics

We use the GADGET-3 code (Springel 2005) where the collisionless dynamics of gravity is solved by a tree-based method (Barnes & Hut 1986), while the hydrodynamics is solved by the smoothed particle hydrodynamics (SPH) method (Gingold & Monaghan 1977; Lucy 1977). We have implemented a modified version of SPH, called SPHGAL, in GADGET-3 which shows a significantly improved numerical accuracy in several aspects (Hu et al. 2014). More specifically, we adopt the pressure-energy formulation of SPH<sup>1</sup> (Read, Hayfield & Agertz 2010; Saitoh & Makino 2013) which is able to properly follow fluid instabilities and mixing without developing severe numerical artefacts commonly found in traditional SPH (e.g. Agertz et al. 2007). The so-called grad-h correction term is included following Hopkins (2013) to ensure the conservation properties when the smoothing length varies significantly (e.g. at strong shocks). The smoothing length is set such that there are  $N_{\text{ngb}} = 100$  particles within a smoothing kernel. We use a Wendland  $C^4$  kernel function that has been shown to be stable against pairing instability for large  $N_{\text{ngb}}$  (Dehnen & Aly 2012), which is necessary for reducing the ‘E<sub>0</sub>-error’ (Read et al. 2010) and therefore improving the numerical convergence rate (Dehnen & Aly 2012; Read & Hayfield 2012). We adopt artificial viscosity to properly model shocks with an efficient switch that only operates at converging flows, similar to the prescriptions presented in Morris & Monaghan (1997) and Cullen & Dehnen (2010). We also include artificial thermal conduction (Price 2008) but only in converging flows to smooth the thermal energy discontinuities, which can lead to severe noise at strong shocks when the pressure-energy formulation is used. The viscosity and conduction coefficients are varied in the range of [0.1,1] and [0,1], respectively. Our SPH scheme shows satisfactory behaviours and accuracies in various idealized numerical tests presented in Hu et al. (2014).

### 2.2 Chemistry model

Our model of the chemistry and cooling follows closely the implementation in the SILCC project (Walch et al. 2015; Girichidis et al. 2016), based on earlier work by Nelson & Langer (1997), Glover & Mac Low (2007) and Glover & Clark (2012b). We track six chemical species:  $H_2$ ,  $H^+$ , CO, H,  $C^+$ , O and free electrons. Only the first three species are followed explicitly, i.e. their abundances are directly integrated based on the rate equations in our chemistry network. The fractional abundance of neutral hydrogen is given by

$$x_{H^0} = 1 - 2x_{H_2} - x_{H^+}, \quad (1)$$

where  $x_i$  denotes the fractional abundance of species  $i$ ; note that all fractional abundances quoted here are relative to the number of H nuclei. Silicon is assumed to be present in singly ionized form (i.e. as  $Si^+$ ) throughout the simulation, while carbon and oxygen may either be present as  $C^+$  and O, or in the form of CO, which leads to

$$\begin{aligned} x_{C^+} &= x_{C,\text{tot}} - x_{CO}, \\ x_O &= x_{O,\text{tot}} - x_{CO}, \end{aligned} \quad (2)$$

<sup>1</sup> We use the pressure-energy SPH instead of the pressure-entropy SPH as used in (Hu et al. 2014) for reasons described in Appendix A.

where  $x_{\text{C,tot}}$  and  $x_{\text{O,tot}}$  are the abundances of the total carbon and oxygen, respectively. Finally, the abundance of free electron is given by

$$x_{e^-} = x_{\text{H}^+} + x_{\text{C}^+} + x_{\text{Si}^+}. \quad (3)$$

A list of our chemical reactions for  $\text{H}_2$  and  $\text{H}^+$  is summarized in table 1 of Micic et al. (2012). The  $\text{H}^+$  is formed via collisional ionization of hydrogen with free electrons and cosmic rays, and is depleted through electron recombination in both the gas phase and on the surfaces of dust grains. The  $\text{H}_2$  is formed on the surfaces of dust grains and is destroyed via ISRF photodissociation, cosmic ray ionization, and collisional dissociation (with  $\text{H}_2$ ,  $\text{H}$  and free electrons).

Our chemical model also includes a treatment of carbon chemistry, following the scheme introduced in Nelson & Langer (1997). It assumes that the rate-limiting step of CO formation is the process  $\text{C}^+ + \text{H}_2 \rightarrow \text{CH}_2^+$ . The  $\text{CH}_2^+$  may either react with atomic oxygen and form CO or be destroyed via ISRF photodissociation. However, we will not go into detailed investigation about CO in this work, as a proper modelling for CO formation is beyond our resolution limit. This is especially so in low-metallicity environments where CO only resides in regions of very high density (Glover & Clark 2012c).

### 2.3 Cooling/heating processes

We include a set of important non-equilibrium cooling and heating processes. The term ‘non-equilibrium’ refers to the fact that the processes depend not only on the local density and temperature of the gas but also on its chemical abundances of species, which may not be in chemical equilibrium. Cooling processes include fine structure lines of  $\text{C}^+$ ,  $\text{O}$  and  $\text{Si}^+$ , the rotational and vibrational lines of  $\text{H}_2$  and CO, the hydrogen Lyman  $\alpha$  line, the collisional dissociation of  $\text{H}_2$ , the collisional ionization of  $\text{H}$ , and the recombination of  $\text{H}^+$  in both the gas phase and on the surfaces of dust grains. Heating processes include photoelectric effects from dust grains and polycyclic aromatic hydrocarbons, ionization by cosmic rays, the photodissociation of  $\text{H}_2$ , the UV pumping of  $\text{H}_2$  and the formation of  $\text{H}_2$ .

We do not follow the non-equilibrium cooling and heating processes in high-temperature regimes. For  $T > 3 \times 10^4$  K, we adopt a cooling function presented in Wiersma, Schaye & Smith (2009), which assumes that the ISM is optically thin and is in ionization equilibrium with a cosmic UV background from Haardt & Madau (2001). The total cooling rate depends on the temperature and density of the gas as well as the abundance of heavy elements. We trace 11 individual elements ( $\text{H}$ ,  $\text{He}$ ,  $\text{C}$ ,  $\text{N}$ ,  $\text{O}$ ,  $\text{Ne}$ ,  $\text{Mg}$ ,  $\text{Si}$ ,  $\text{S}$ ,  $\text{Ca}$  and  $\text{Fe}$ ) for both gas and star particles based on the implementation in Aumer et al. (2013).

### 2.4 Shielding of the ISRF

Shielding from the ISRF affects both the chemistry and the cooling/heating processes. For the hydrogen chemistry, the  $\text{H}_2$  ISRF photodissociation rate,  $R_{\text{pd,H}_2}$ , is attenuated by both the dust shielding and the  $\text{H}_2$  self-shielding:

$$R_{\text{pd,H}_2} = f_{\text{sh}} R_{\text{pd,H}_2,\text{thin}}, \quad (4)$$

where  $R_{\text{pd,H}_2,\text{thin}} = 3.3 \times 10^{-11} G_0 \text{ s}^{-1}$  is the unattenuated photodissociation rate,  $G_0$  is the strength of the ISRF relative to the solar neighbourhood value estimated by Habing (1968). The total attenuation factor is  $f_{\text{sh}} = f_{\text{dust,H}_2} f_{\text{self,H}_2}$  where  $f_{\text{dust,H}_2}$  and  $f_{\text{self,H}_2}$  are

the attenuation factors by dust extinction and by  $\text{H}_2$  self-shielding, respectively. We adopt

$$f_{\text{dust,H}_2} = \exp(-D\sigma_{\text{dust}}N_{\text{H,tot}}), \quad (5)$$

where  $D$  is the DGR relative to the Milky Way value ( $\sim 1$  per cent) and  $\sigma_{\text{dust}} = 2 \times 10^{-21} \text{ cm}^2$  is the averaged cross-section of dust extinction. The  $\text{H}_2$  self-shielding is related to the  $\text{H}_2$  column density  $N_{\text{H}_2}$  using the relation given by Draine & Bertoldi (1996). A similar treatment is used for the carbon chemistry. The CO photodissociation rate is attenuated by the dust extinction,  $\text{H}_2$  shielding and CO self-shielding. A more detailed description can be found in Walch et al. (2015).

Dust extinction also reduces the photoelectric heating rate by blocking the radiation in the energy range between 6 and 13.6 eV. We adopt the heating rate given by Bakes & Tielens (1994) and Bergin et al. (2004) of

$$\Gamma_{\text{pe}} = 1.3 \times 10^{-24} \epsilon D G_{\text{eff}} n \text{ erg s}^{-1} \text{ cm}^{-3}, \quad (6)$$

where  $G_{\text{eff}} = G_0 \exp(-1.33 \times 10^{-21} DN_{\text{H,tot}})$  is the attenuated radiation strength,  $n$  is the number density of the gas and  $\epsilon$  is the heating efficiency defined as

$$\epsilon = \frac{0.049}{1 + (0.004 \psi)^{0.73}} + \frac{0.037(T/10\,000)^{0.7}}{1 + 2 \times 10^{-4} \psi}, \quad (7)$$

where  $\psi = G_{\text{eff}} T^{0.5} / n_{e^-}$  and  $n_{e^-}$  is the electron number density.

To calculate the column densities relevant for shielding we have incorporated the TREECOL algorithm (Clark, Glover & Klessen 2012) into our version of GADGET-3. Unlike the extragalactic UV background that is external to the simulated galaxy, the sources of the ISRF are the young stars embedded in the ISM. This means that the column densities should not be integrated over the entire disc, but have to be truncated at certain length scales. Ideally, one should integrate the column densities up to individual stars that contribute to the local radiation field separately, but this would entail performing a computationally expensive radiative transfer calculation on every timestep, which is impractical. Instead, we make the simplifying assumption that the material beyond a predefined shielding length  $L_{\text{sh}}$  from a local gas particle is not relevant for the shielding (see e.g. Dobbs et al. 2008 or Smith et al. 2014, who make similar approximations in their galactic simulations). We justify this assumption by noting that for gas particles in dense clouds – the only ones which are significantly shielded – the dominant contribution to the shielding generally comes from gas and dust in the cloud itself or in its immediate vicinity, rather than from the diffuse gas between the clouds. Therefore, provided we adopt a value for  $L_{\text{sh}}$  that is larger than the typical size of the dense clouds, our results should be insensitive to the precise value chosen for  $L_{\text{sh}}$ .<sup>2</sup>

For each gas particle, TREECOL defines  $N_{\text{pix}}$  equal-area pixels using the HEALPIX algorithm (Górski & Hivon 2011) and computes the column density within each pixel out to  $L_{\text{sh}}$ . We set  $N_{\text{pix}} = 12$  and  $L_{\text{sh}} = 50$  pc throughout this work. However, we explore in Appendix C2 the effect of varying  $L_{\text{sh}}$  within the range 20–200 pc and show that as expected our results are insensitive to the precise value of  $L_{\text{sh}}$  within this range.

<sup>2</sup> Along a few lines of sight which intercept other dense clouds, taking a small  $L_{\text{sh}}$  causes us to potentially underestimate the shielding. However, the local radiation field will always be dominated by other lines of sight with low column density.

## 2.5 Star formation model

Unlike cloud-scale simulations, our mass resolution is not sufficient to follow the gravitational collapse of the gas to densities where it will inevitably end up in stars. Instead, we define an instantaneous SFR for each gas particle to estimate how much gas is expected to be converted into stars:  $\text{SFR}_{\text{gas}} = \epsilon_{\text{sf}} m_{\text{gas}} / t_{\text{ff}}$ , where  $t_{\text{ff}} = (4\pi G \rho_{\text{gas}})^{-0.5}$  is the local free-fall time,  $\epsilon_{\text{sf}}$  is the star formation efficiency, and  $\rho_{\text{gas}}$  is the volumetric density of gas. We set  $\epsilon_{\text{sf}} = 0.02$  to account for the inefficient star formation which might originate from the sub-resolution turbulent motions (Krumholz & McKee 2005). We assume the gas is ‘star-forming’ ( $\text{SFR}_{\text{gas}} > 0$ ) only if the gas has  $n_{\text{H}} \geq n_{\text{H, th}}$ ,  $T \leq T_{\text{th}}$  and negative velocity divergence. We choose  $n_{\text{H, th}} = 100 \text{ cm}^{-3}$  as this is the typical densities of the giant molecular clouds in our Galaxy, the reservoir gas for star formation for which  $\epsilon_{\text{sf}} = 0.02$  is defined. We also set  $T_{\text{th}} = 100 \text{ K}$  to ensure that we do not attempt to form stars in hot dense gas which has a high Jeans mass. In practice, most gas with  $n_{\text{H}} > n_{\text{H, th}}$  would actually be colder than 100 K (cf. Fig. 9). Our definition of ‘star-forming gas’ is very simplistic. In reality, stars may still form out of a gas cloud with  $n_{\text{H}} < n_{\text{H, th}}$  if the sub-resolution density structure is very clumpy. Note that the appropriate choice of  $n_{\text{H, th}}$  and  $\epsilon$  also depends on resolution. With higher resolutions, one should be able to follow the gravitational collapse to smaller scales and denser environments (e.g. individual molecular cores), and therefore both  $n_{\text{H, th}}$  and  $\epsilon$  should be set higher.

We adopt the commonly used stochastic approach of star formation: in each timestep of size  $\Delta t$ , a star-forming gas particle has a probability of  $\epsilon_{\text{sf}} \Delta t / t_{\text{ff}}$  to be converted into a star particle of the same mass. Since  $\Delta t \ll t_{\text{ff}} / \epsilon_{\text{sf}}$  almost always holds, the conversion test in each timestep is a rare event Bernoulli trial and the number of ‘success’ events during a time period of  $t_{\text{ff}} / \epsilon_{\text{sf}}$  for a given  $\text{SFR}_{\text{gas}}$  follows a Poisson distribution with the parameter  $\lambda = 1$  (which is the expectation value), i.e. one gas particle would be converted into a star particle in a time period of  $t_{\text{ff}} / \epsilon_{\text{sf}}$  on average. Note that the ratio of the standard deviation to the expectation value for a Poisson distribution is  $\lambda^{-0.5}$ , so the actual star formation time-scale can deviate from the input by  $\approx 100$  per cent if we look at a single particle. Only when we look at the averaged SFR of a group of particles would the random fluctuation be reduced to a satisfying level.

The instantaneous SFR of gas particles, however, is not an observable. In fact, it is merely an estimate for the star formation that will happen in the next timesteps. Therefore, we measure the SFR in a given region by the total mass of newly formed star particles with age less than  $t_{\text{SF}}$  in that region divided by  $t_{\text{SF}}$ , which is set to be 5 Myr in this work. Such a definition is more compatible with what is measured in observations than the instantaneous SFR assigned to the gas particles based on the adopted star formation model. These two definitions of SFR give almost identical results if we sum over a large enough region (e.g. the total SFR of the galaxy), although locally (both in space and in time) they can be quite different (cf. Section 4.7.1). Throughout this work we will adopt this definition when we present the SFR (except for Fig. 16 where both definitions are shown).

## 2.6 Stellar feedback and metal enrichment

### 2.6.1 Supernova Type II (SNII)

We assume each star particle represents a stellar population of mass  $m_{\text{star}}$  and calculate the corresponding mass  $\delta m_{\text{SNII}}$  that will end up in SNII. For our adopted Kroupa initial mass function (Kroupa 2001),

we have  $\delta m_{\text{SNII}} \simeq 0.12 m_{\text{star}}$ . When the age of a star particle reaches 3 Myr, we return  $\delta m_{\text{SNII}}$  of mass to the ISM with enriched metal abundances according to the metallicity dependent yields given by Woosley & Weaver (1995). The returned mass is added to the nearest 100 neighbouring gas particles weighted by the smoothing kernel. The energy budget for a  $m_{\text{star}}$  stellar population is  $N_{\text{SNII}} E_{51}$  where  $E_{51} = 10^{51} \text{ erg}$  is the typical energy for a single SN event and  $N_{\text{SNII}} = m_{\text{star}} / 100 M_{\odot}$  is the number of SN events.

Physically, a supernova remnant (SNR) should first go through a non-radiative phase where momentum is generated (the Sedov phase) until radiative cooling kicks in and the total momentum gradually saturates (e.g. Blondin et al. 1998; Ostriker & McKee 1988). However, numerically, the resolution requirement for modelling the correct evolution of an SNR is very demanding (see e.g. Walch & Naab 2015). It has long been recognized that insufficient resolution leads to numerical overcooling: most of the injected energy is radiated away before it can significantly affect the ISM. As shown in Dalla Vecchia & Schaye (2012), this occurs when the cooling time is much shorter than the response time of the gas for a given resolution. For usual implementations in SPH simulations where  $m_{\text{star}} = m_{\text{gas}}$ , this means that the returned mass  $\delta m_{\text{SNII}}$  is always much smaller than  $m_{\text{gas}}$  and the SNR would be poorly resolved. The situation would not be alleviated by reducing  $m_{\text{gas}}$  as long as  $m_{\text{star}} = m_{\text{gas}}$  is assumed. Dalla Vecchia & Schaye (2012) circumvent this issue by injecting the energy to fewer neighbouring particles and, if necessary, stochastic injection which groups several SN events into a single energetic one. By doing so they guarantee that the gas would always be heated to the local minimum of the cooling function and has more time to develop the blast wave.

Although grouping several SN events into one energetic explosion enhances the dynamical impact of SN feedback on the ISM, it also coarsens the granularity (both spatial and temporary) of the SN events. As shown in Kim & Ostriker (2015), a single energetic explosion overproduces both the momentum and energy compared to a series of spatially coherent explosions with the same amount of total energy. The difference would probably be even more severe if the explosions occur at different locations. Therefore, for a given energy budget, coarser SN sampling can overestimate the impact of SN feedback. With the SN sampling as a free parameter, the effect of feedback becomes tunable or even arbitrary. Indeed, in large-scale cosmological simulations with necessarily compromised resolutions, the feedback has to be calibrated by fitting to observations (Schaye et al. 2015) and thus can only be regarded as a phenomenological model. However, if one’s goal is to directly resolve individual blast waves without tunable parameters, as we try to do in this work, then the SN sampling should be set to the physical one by making sure that each SN event has the canonical energy of  $10^{51} \text{ erg}$ .

As the star particle mass reaches  $m_{\text{star}} < 100 M_{\odot}$ , the energy budget for a star particle becomes smaller than  $E_{51}$  (i.e.  $N_{\text{SNII}} < 1$ ), which is also unphysical. In this work we adopt an ansatz similar to the stochastic injection in Dalla Vecchia & Schaye (2012). We let star particles with the appropriate age explode as SNII with a probability of  $m_{\text{gas}} / 100 M_{\odot}$  and with an energy of  $E_{51}$ . In this work,  $m_{\text{gas}} = 4 M_{\odot}$  and therefore each star particle has a 4 per cent chance of producing a Type II SN. Our resolution is close to the  $m_{\text{gas}} = 1 M_{\odot}$  requirement for a reasonably converged energy and momentum evolution of an SNR as shown in the resolution study in Appendix B.

Note that despite the fact that our particle mass ( $4 M_{\odot}$ ) is comparable to the mass of a single star, the star particles should still be considered as stellar populations instead of individual stars. A

star particle, once formed, is only a representative point of the collisionless distribution function that describes the stellar system of the galaxy. With this interpretation, there is no conceptual issue even when the particle mass reaches sub-solar scales, though the system might be overresolved. The collisional dynamics of star clusters is suppressed by our gravitational softening (2 pc), and thus could only be included in a sub-resolution model, which is not considered in this work.

### 2.6.2 Supernova Type Ia (SNIa) and asymptotic giant branch (AGB) stars

We include feedback by SNIa and AGB stars based on the implementation presented in Aumer et al. (2013). For SNIa, we adopt a delay time distribution (DTD), the SN rate as a function of time for a given stellar population formed in a single burst. The DTD has a power-law shape  $\sim t^{-1}$  where  $t$  is the stellar age, with the normalization of two SNIa events per 1000  $M_{\odot}$  (Maoz & Mannucci 2012). The amount of mass returned to the ISM is calculated by sampling the DTD with a 50 Myr time bin, and the metal yields based on Iwamoto et al. (1999). Similarly, the mass returned by the AGB stars is calculated from the metal yields presented in Karakas (2010) with the same time bin as SNIa. Assuming an outflow velocity  $v = 3000 \text{ km s}^{-1}$  and  $10 \text{ km s}^{-1}$  for SNIa and AGB stars, respectively, we return energy of  $0.5 \delta m v^2$  into the ISM where  $\delta m$  is the returned mass in a given time bin, though in our simulations their effect is sub-dominant compared to the SNII feedback.

### 2.7 Timestep limiter

We include a timestep limiter similar to Saitoh & Makino (2009) and Durier & Dalla Vecchia (2012) to correctly model the strong shocks. For each active particle  $i$ , we identify any neighbouring particles within its smoothing kernel whose timesteps are four times longer than the timestep of particle  $i$  and force them to be active at the next synchronization point. In addition, we re-calculate the hydrodynamical properties for the particles which we inject the feedback energy into and make them active such that their timesteps will be assigned correctly at the next synchronization point. Note that by modifying a particle's timestep before it completes its full kick-drift-kick procedure we necessarily violate energy conservation, but to a much lesser extent than a numerical scheme without the timestep limiter.

### 2.8 Numerical resolution

The mass resolution of an SPH simulation depends not only on the particle mass ( $m_{\text{gas}}$ ) but also the number of particles within a smoothing kernel ( $N_{\text{ngb}}$ ). For a given kernel function and  $m_{\text{gas}}$ , using more particles in a kernel means smoothing over more mass and hence worse resolution. However, because of the low-order nature of SPH, a relatively large  $N_{\text{ngb}}$  is required to reduce the so-called  $E_0$ -error, a zeroth order numerical noise induced by particle disorder (see e.g. Read et al. 2010; Dehnen & Aly 2012). We adopt  $N_{\text{ngb}} = 100$  as a compromise between suffering too much from the  $E_0$ -error ( $N_{\text{ngb}}$  too small) and oversmoothing ( $N_{\text{ngb}}$  too large). It seems reasonable to regard the kernel mass  $M_{\text{ker}} \equiv N_{\text{ngb}} m_{\text{gas}}$  as the resolution of SPH simulations. However, different kernel functions entail different extent of smoothing and different scales of compact support ( $H$ ). A Gaussian kernel is an extreme example with infinite  $N_{\text{ngb}}$  and  $H$  while its smoothing scale is obviously finite. The same

$N_{\text{ngb}}$  (and hence  $H$ ) can therefore mean very different resolutions depending on the adopted kernel. A more physical meaningful way is required to define a length scale which reflects the true extent of smoothing. Dehnen & Aly (2012) proposed to define the smoothing scale as  $h = 2\sigma$ , where  $\sigma$  is the second moment of the kernel function (or 'standard deviation'). Following such a definition,  $h \approx 0.55 H$  for the commonly used cubic spline kernel and  $h \approx 0.45 H$  for our adopted Wendland  $C^4$  kernel. Therefore, the mass resolution would be  $N_{\text{ngb}} m_{\text{gas}} (h/H)^3 \approx 0.1 M_{\text{ker}}$ . For  $m_{\text{gas}} = 4 M_{\odot}$  in our simulations (see Section 4.1), this means that  $40 M_{\odot}$  is the mass scale for which we define the local density of a particle. Everything below  $40 M_{\odot}$  is blurred by smoothing.

#### 2.8.1 Jeans mass criterion

In hydrodynamical simulations that include self-gravity (as we do in this work), one important scale to resolve is the Jeans mass such that

$$M_J \geq N_J M_{\text{ker}}, \quad (8)$$

where  $M_J$  is the Jeans mass and  $N_J$  is a prescribed number to ensure  $M_J$  is well resolved (Bate & Burkert 1997; Robertson & Kravtsov 2008). When the Jeans mass is not resolved ( $M_J < N_J M_{\text{ker}}$ ), perturbations on all scales down to the resolution limit ( $\sim M_{\text{ker}}$ ) should collapse physically. However, perturbations can be created by numerical noise, triggering gravitational collapse, which tends to occur when the perturbations are only marginally resolved. Equation (8) makes sure that perturbations near the resolution limit are physically stable and all perturbations that collapse are well resolved and are of physical origin rather than numerical noise. The choice of  $N_J$  is somewhat arbitrary, as it is difficult to judge whether a collapse is physical or numerical. Commonly suggested values in the literature are in the range of  $N_J = 4-15$  (e.g. Robertson & Kravtsov 2008; Hopkins, Quataert & Murray 2011; Richings & Schaye 2016), which seems to be quite stringent considering the resolution is about  $0.1 M_{\text{ker}}$ . This may be related to the smaller  $N_{\text{ngb}}$  commonly used which is more prone to noise. As such, we take  $N_J = 1$ , which means that we require one kernel mass to resolve the Jeans mass. According to the gas distribution in the phase diagram (see Fig. 9), our maximum number density satisfying  $M_J \geq M_{\text{ker}}$  is about  $200 \text{ cm}^{-3}$ , which corresponds to the smoothing length  $H \approx 2 \text{ pc}$ . This motivates us to choose our gravitational softening length to be also 2 pc.

A commonly adopted approach to ensure that the Jeans mass is formally resolved throughout the simulations is to introduce a 'pressure floor' which artificially stabilizes the gas (by adding pressure by hand) when it becomes Jeans-unresolved (violating equation 8; e.g. Robertson & Kravtsov 2008; Hopkins et al. 2011; Renaud et al. 2013; Vogelsberger et al. 2014; Schaye et al. 2015). However, physically, a Jeans-unresolved perturbation should collapse. In fact, the only way to properly follow the evolution of the Jeans-unresolved gas is to increase the resolution so that it becomes Jeans-resolved. When a gas cloud becomes Jeans-unresolved, the credibility of gravitational effect is lost irrespective of whether a pressure floor (or any other similar approach that makes the equation of state artificially stiffer) is used or not. Besides, the pressure floor makes the gas artificially stiff and thus also sabotages the accuracy of hydrodynamical effects. We therefore refrain from using a pressure floor and simply try to keep the majority of gas Jeans-resolved by using high enough resolution.

### 3 THE ISM IN EQUILIBRIUM

Before delving into the non-equilibrium evolution of the ISM and the complicated interplay of all the physical processes, in this section we investigate the ISM properties in both thermal and chemical equilibrium under typical conditions in dwarf galaxies. To see how a system evolves towards chemical and thermal equilibrium in a uniform and static medium, we run the code with only the cooling and chemistry modules while turning the gravity and SPH solvers off. The density and column density are directly specified as input parameters for each particle rather than calculated with SPHGAL or TREECOL. The initial temperature is set to be  $10^4$  K. We set the metallicity  $Z = 0.1 Z_\odot$  and the cosmic ray ionization rate  $\zeta = 3 \times 10^{-18} \text{ s}^{-1}$ .

The parameters that directly affect  $H_2$  formation and destruction are  $G_0$  and  $D$  (equations 10 and 11). The strength of the ISRF in the diffuse ISM is expected to correlate with the SFR density. As will be shown in Section 4, the total SFR of our simulated galaxy is  $\approx 10^{-3} M_\odot \text{ yr}^{-1}$  and the radius of the star-forming region is  $\approx 2$  kpc. Thus, the SFR surface density is  $\approx 8 \times 10^{-5} M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$ , which is about 10 times smaller than the value in the solar neighbourhood. Assuming the correlation between SFR surface density and  $G_0$  is linear (e.g. Ostriker, McKee & Leroy 2010),  $G_0 = 0.17$  would be a plausible choice. On the other hand, the emergent radiation from the less dusty star-forming regions in low-metallicity environments seems to be stronger than in metal-rich ones due to the higher escape fraction of UV photons (see e.g. Bolatto et al. 2011; Cormier et al. 2015). Therefore, the resulting  $G_0$  in dwarf galaxies might be as strong as that in the solar neighbourhood. The observed DGR of a galaxy has been shown to scale linearly with its metallicity (though with significant scatter). This implies  $D = 0.1$  for our adopted metallicity  $Z = 0.1 Z_\odot$ . However, below  $Z \approx 0.2 Z_\odot$ , the DGR starts to fall below the linear DGR– $Z$  relation (Rémy-Ruyer et al. 2014).

In this work we explore three different combinations of  $G_0$  and  $D$ . We will use the following naming convention.

- (i) *GID01*:  $G_0 = 1.7, D = 0.1$ ,
- (ii) *GID001*:  $G_0 = 1.7, D = 0.01$ ,
- (iii) *G01D01*:  $G_0 = 0.17, D = 0.1$ .

#### 3.1 Chemical equilibrium

##### 3.1.1 The $H_2$ formation/destruction time-scale

We adopt the  $H_2$  formation rate:

$$\dot{n}_{H_2} = R_{\text{form}} D n_H n_{H_1} \quad (9)$$

where  $R_{\text{form}}$  is the rate coefficient (Hollenbach & McKee 1979), and  $n_{H_2}$ ,  $n_{H_1}$  and  $n_H$  are the number densities of molecular hydrogen, atomic hydrogen, and hydrogen nuclei, respectively.  $D$  is the DGR normalized to the Milky Way value such that  $D = 1$  means the DGR = 1 per cent. At low temperatures ( $T \lesssim 100$  K), the rate coefficient  $R_{\text{form}} \approx 3 \times 10^{-17} \text{ cm}^3 \text{ s}^{-1}$  is relatively insensitive to temperature variations, and the solution to equation (9) is  $n_{H_2}/n_H = 1 - \exp(-t/t_F)$  with the formation time-scale:

$$t_F = (R_{\text{form}} D n_H)^{-1} \approx \frac{1}{n_1 D} \text{ Gyr}, \quad (10)$$

where  $n_1 \equiv n_H/1 \text{ cm}^{-3}$ .

Meanwhile, the dominant process of  $H_2$  destruction is photodissociation by the ISRF in the Lyman and Werner bands (11.2–13.6 eV). From equation (4), the destruction time-scale is

$$t_D = (R_{\text{pd},H_2})^{-1} \approx \frac{1}{f_{\text{sh}} G_0} \text{ kyr}. \quad (11)$$

Therefore,  $H_2$  clouds exposed to an unattenuated radiation field will be quickly destroyed, independent of the gas density. However, the photodissociation itself is also an absorption process, which attenuates the UV radiation efficiently and thus lengthens  $t_D$  as the  $H_2$  column density accumulates ( $H_2$  self-shielding).

To illustrate the formation and destruction of  $H_2$  for different ISM parameters, we set up hydrogen column densities from  $N_{H,\text{min}} = 10^{16} \text{ cm}^{-2}$  to  $N_{H,\text{max}} = 10^{24} \text{ cm}^{-2}$  sampled by 32 768 points equally spaced in  $\log_{10} N_H$  and evolve the chemistry network including cooling and heating. This corresponds to a one-dimensional absorbing slab in a homogeneous medium. The  $H_2$  column density  $N_{H_2}$  is obtained by direct integration:

$$N_{H_2}(N_H) = \int_{N_{H,\text{min}}}^{N_H} x_{H_2}(N'_H) dN'_H. \quad (12)$$

We define the  $H_2$  mass fraction  $f_{H_2} \equiv 2x_{H_2} = 2n_{H_2}/n_H$  and therefore  $f_{H_2} = 1$  means that the hydrogen is fully molecular.

In the upper three rows in Fig. 1 we show the time evolution of  $f_{H_2}$  in a uniform medium of  $n_H = 100 \text{ cm}^{-3}$  in *GID01*, *GID001* and *G01D01* from top to bottom. Initially, the medium is totally atomic ( $f_{H_2} = 0$ ) in the left-hand panels and totally molecular ( $f_{H_2} = 1$ ) in the right-hand panels. The formation times are  $t_F = 100, 1000$  and  $100$  Myr for *GID01*, *GID001* and *G01D01*, respectively. If the medium is initially atomic, in a few  $t_F$ , gas at sufficiently high  $N_H$  becomes totally molecular and the system reaches chemical equilibrium, which is consistent with equation (10). If the initial medium is molecular, the gas would be photodissociated starting from the surface (low  $N_H$ ). In the unattenuated region, equation (11) suggests rapid destruction on time-scales of kyr, much shorter than  $t_F$ . However, only the thin surface directly exposed to the UV radiation would be destroyed very promptly. Due to  $H_2$  self-shielding, the dissociating front propagates slowly into high  $N_H$  and the system reaches chemical equilibrium on time-scales much longer than 1 kyr. On the other hand, the free-fall time at  $n_H = 100 \text{ cm}^{-3}$  is  $t_{\text{ff}} \approx 5$  Myr. Therefore, in the ISM of dwarf galaxies, the time-scale to reach equilibrium, either starting from an atomic or molecular medium, is quite long compared to the local free-fall time.

##### 3.1.2 The $H_2$ fraction in chemical equilibrium

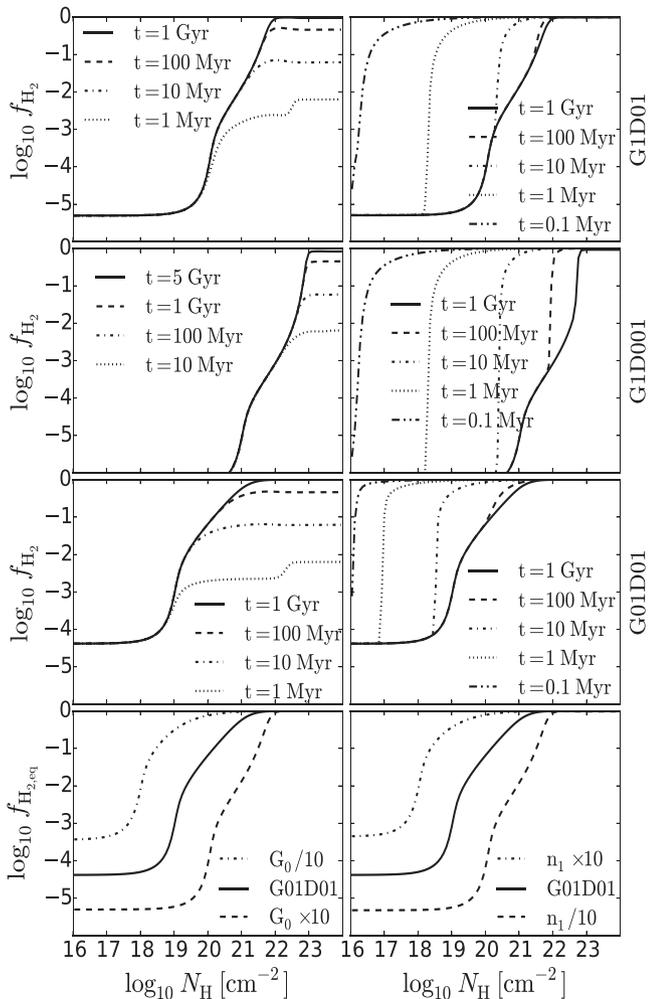
An unperturbed cloud will eventually reach a steady state where the  $H_2$  formation rate equals its destruction rate:

$$\dot{n}_{H_2} = R_{\text{form}} D n_H n_{H_1} - R_{\text{pd},H_2} n_{H_2} = 0. \quad (13)$$

Together with  $f_{H_2} = 2n_{H_2}/n_H$  and  $n_H = n_{H_1} + 2n_{H_2}$  (assuming the ionization fraction of hydrogen is zero), the ratio of  $H_2$  to  $H_1$  at chemical equilibrium can be written as

$$r_{H_2/H_1} \equiv \frac{f_{H_2}}{1 - f_{H_2}} = 1.8 \times 10^{-6} \frac{n_1 D}{f_{\text{sh}} G_0}. \quad (14)$$

In unshielded regions,  $f_{H_2} \approx r_{H_2/H_1}$  tends to be very low for typical conditions. Only in well-shielded environments where very few  $H_2$ -dissociating photons are present can  $f_{H_2}$  grow to a significant value. The  $H_2$  profile is dictated by the value of  $n_1 D/f_{\text{sh}} G_0$ , which can be viewed as a dimensionless parameter representing the capability of  $H_2$  formation relative to its destruction. In the bottom row of Fig. 1,

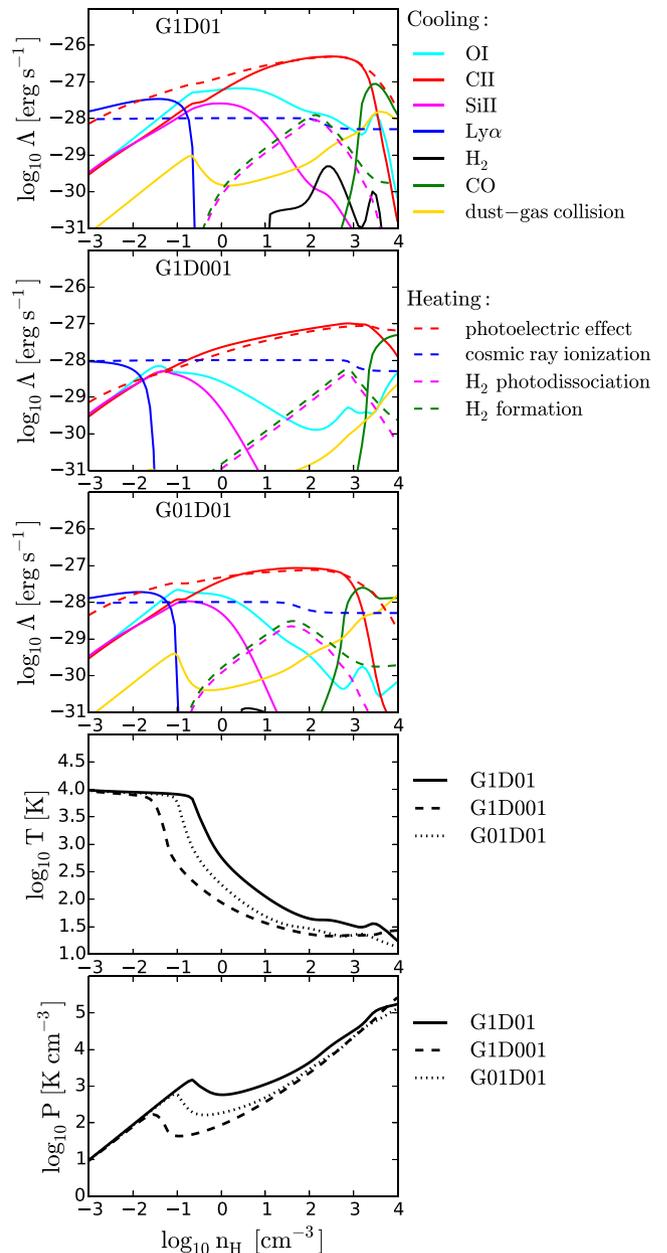


**Figure 1.** Upper three rows: time evolution of  $f_{\text{H}_2}$  in a uniform medium of  $n_{\text{H}} = 100 \text{ cm}^{-3}$  in *G1D01* (first row), *G1D001* (second row) and *G01D01* (third row). The medium is initially totally atomic ( $f_{\text{H}_2} = 0$ ) in the left-hand panels and totally molecular ( $f_{\text{H}_2} = 1$ ) in the right-hand panels. Bottom row: the equilibrium  $\text{H}_2$  profile ( $f_{\text{H}_2}$  versus  $N_{\text{H}}$ ) with different  $G_0$  (left-hand panel) and  $n_1$  (right-hand panel). For a given  $D$ , the equilibrium  $\text{H}_2$  profile is self-similar depending on the dimensionless parameter  $n_1/G_0$ .

we show the equilibrium  $\text{H}_2$  profile ( $f_{\text{H}_2}$  versus  $N_{\text{H}}$ ) with different  $G_0$  and  $n_1$ . For a given value of  $D$ , increasing (decreasing)  $G_0$  by a factor of 10 has the same effect as decreasing (increasing)  $n_1$  by the same factor. The equilibrium  $\text{H}_2$  profile is thus a self-similar solution depending on the value of  $n_1/G_0$  (Sternberg 1988; McKee & Krumholz 2010; Sternberg et al. 2014). Note that although the equilibrium profile is self-similar, the time required for reaching the equilibrium ( $t_{\text{F}}$  or  $t_{\text{D}}$ ) is not.

### 3.2 Thermal equilibrium

We set up the hydrogen number density  $n_{\text{H}}$  from  $10^{-4}$  to  $10^4 \text{ cm}^{-3}$  sampled by 32 768 points equally spaced in  $\log_{10} n_{\text{H}}$  and let the system cool from  $10^4 \text{ K}$  until it reaches thermal equilibrium. Here we assume the column density to be  $N_{\text{H}} = n_{\text{H}} H$  where  $H$  is the SPH smoothing length. The  $\text{H}_2$  column density is obtained by the approximation  $N_{\text{H}_2} = x_{\text{H}_2} N_{\text{H}}$ . This is the minimum column density that would be assigned to a gas particle for the given density in our simulations if there were no other contributing material integrated



**Figure 2.** The cooling and heating rates for gas in thermal equilibrium in *G1D01* (first panel), *G1D001* (second panel) and *G01D01* (third panel). The fourth panel shows the equilibrium temperature versus density, and the fifth panel shows the equilibrium pressure versus density. The maximum density with an equilibrium temperature about  $10^4 \text{ K}$  is  $n_{\text{cool}} = 0.25, 0.03$  and  $0.1 \text{ cm}^{-3}$  in *G1D01*, *G1D001* and *G01D01*, respectively.

by TREECOL. The smoothing length is calculated assuming the same mass resolution ( $m_{\text{gas}} = 4 M_{\odot}$ ,  $N_{\text{ngb}} = 100$ ) as in our simulations in Section 4.

In the upper three panels in Fig. 2, we show the individual cooling and heating processes in thermal equilibrium in *G1D01*, *G1D001* and *G01D01*. The fourth panel shows the equilibrium temperature versus density, and the fifth panel shows the equilibrium pressure versus density. For almost the entire range of density, the photoelectric effect is the main heating mechanism. In *G1D001*, the photoelectric effect becomes less efficient (due to the low DGR), and thus cosmic ray ionization dominates at  $n_{\text{H}} < 10^{-1} \text{ cm}^{-3}$ .

The dust–gas collision rate rises with  $n_H$  but only becomes an important cooling mechanism at the highest densities which rarely occur in our simulations. The most dominant coolant in the range of  $n_H = 10^{-1}$ – $10^3 \text{ cm}^{-3}$  is the  $C^+$  fine structure line emission. Below  $10^{-1} \text{ cm}^{-3}$ , as the temperature increases, the O I fine structure emission and the hydrogen Lyman  $\alpha$  emission start to dominate over  $C^+$ . Above  $10^3 \text{ cm}^{-3}$  the gas becomes optically thick to the UV radiation and so the photoelectric heating rate decreases. Meanwhile, the  $C^+$  cooling is taken over by the CO cooling as most carbon is in the form of CO in this regime. The  $H_2$  cooling is unimportant in all cases as it is typically only abundant at densities where the temperature is too low to excite  $H_2$ . In *G1D001* and *G01D01*, the photoelectric heating is less efficient than in *G1D01* due to lower DGR and weaker ISRF, respectively (cf. equation 6). Therefore, the equilibrium temperatures are the highest in *G1D01* (bottom panel of Fig. 2). Note that in low-metallicity and low-density regimes, the time required to reach thermal equilibrium can become quite long (Krumholz 2012; Glover & Clark 2014). Therefore, the gas is expected to be out of thermal equilibrium at low densities (cf. Fig. 9).

The maximum density with an equilibrium temperature about  $10^4 \text{ K}$  is  $n_{\text{cool}} = 0.25, 0.03$  and  $0.1 \text{ cm}^{-3}$  in *G1D01*, *G1D001* and *G01D01*, respectively. As will be shown in Fig. 13,  $n_{\text{cool}}$  determines the maximum radius for star formation as it marks the onset of thermal-gravitational instability (Schaye 2004).

## 4 SIMULATIONS

### 4.1 Initial conditions

We set up the initial conditions using the method developed in Springel (2005). The dark matter halo has a virial radius  $R_{\text{vir}} = 44 \text{ kpc}$  and a virial mass  $M_{\text{vir}} = 2 \times 10^{10} M_{\odot}$ , and follows a Hernquist profile with an NFW-equivalent (Navarro, Frenk & White 1997) concentration parameter  $c = 10$ . The spin parameter for the dark matter halo is  $\lambda = 0.03$ . A disc comprised of both gas and stellar components is embedded in the dark matter halo. The total mass of the disc is  $6 \times 10^7 M_{\odot}$  with a gas fraction of 66 per cent. The small baryonic mass fraction (0.3 per cent) is motivated by the results of abundance matching (Moster et al. 2010; Moster, Naab & White 2013). The disc follows an exponential profile with scalelength of 0.73 kpc. The scalelength is determined by assuming the disc is rotationally supported and the total angular momentum of the disc is 0.3 per cent of that of the dark matter halo, which leads to a one-to-one relation between  $\lambda$  and the scalelength. Since dwarf galaxies are usually expected to have a relatively thick disc (e.g. Elmegreen & Hunter 2015), we set the scaleheight of the disc to be 0.35 kpc. The initial metallicity (gas and stars) is  $0.1 Z_{\odot}$  uniformly throughout the disc with the relative abundances of the various metals the same as in solar metallicity gas. The cosmic ray ionization rate is  $3 \times 10^{-18} \text{ s}^{-1}$ . The particle mass is  $m_{\text{dm}} = 10^4 M_{\odot}$  for the dark matter and  $m_{\text{gas}} = m_{\text{disc}} = 4 M_{\odot}$  for the baryons (both stars and gas). The gravitational softening length is 62 pc for dark matter and 2 pc for baryons. As discussed in Section 3, we explore three different combinations of  $G_0$  and  $D$  as shown in Table 1. Run *G1D01\_noFB* has the same  $G_0$  and  $D$  as *G1D01* while the stellar feedback is switched off.

We use cylindrical coordinates  $R$  and  $z$  to describe the simulations, where  $R$  is the galactocentric radius and  $z$  is the rotation axis of the disc. The origin is chosen at the centre of mass of the stellar disc, so that  $z = 0$  is the mid-plane of the disc. We also use the

**Table 1.** Simulation runs and the corresponding setup.

Name	$G_0$	$D$	Feedback
<i>G1D01</i>	1.7	0.1	Yes
<i>G1D001</i>	1.7	0.01	Yes
<i>G01D01</i>	0.17	0.1	Yes
<i>G1D01_noFB</i>	1.7	0.1	No

spherical coordinate  $r$  to indicate the distance from the origin to a certain radius.

### 4.2 Morphology

In Fig. 3 we show, from left to right, the column density maps of  $H\text{I}$  (first column),  $H_2$  (second column) and  $H^+$  (third column) and temperature maps (slices, fourth column) at  $t = 500 \text{ Myr}$ , where  $t$  is the simulation time. The upper two rows are the face-on and edge-on views for run *G1D01*, while the lower two rows are for run *G1D001*. Fig. 4 shows the same maps as Fig. 3 but for run *G01D01* (upper two rows) and run *G1D01\_noFB* (lower two rows). The temperature maps are slices across the origin so that the face-on slices are at the mid-plane  $z = 0$ . We use different ranges of the colour bars for different chemical species for display purposes.

The ISM is strongly dominated by  $H\text{I}$  in all cases. It has clumpy density structures with many SN-driven bubbles. Most  $H_2$  resides in a thin layer in the mid-plane with little  $H_2$  at large  $|z|$ . The  $H_2$  approximately traces dense gas. The  $H^+$  traces the shells of the SN-driven bubbles and is not confined to the mid-plane. The warm (green) gas occupies most of the volume in almost all cases. The hot (red) gas also fills up a non-negligible amount of volume (mainly in the SN bubbles) while the cold (blue) gas occupies very little volume.

Comparing the runs with feedback,  $H_2$  is most abundant in *G01D01* while almost non-existent in *G1D001*. This is expected since the strong ISRF and low DGR in *G1D001* is a hostile environment for  $H_2$  formation. If we turn off feedback,  $H_2$  becomes more abundant than in *G01D01* (cf. Fig. 6, although it occupies little volume because the gas collapses into massive clumps).

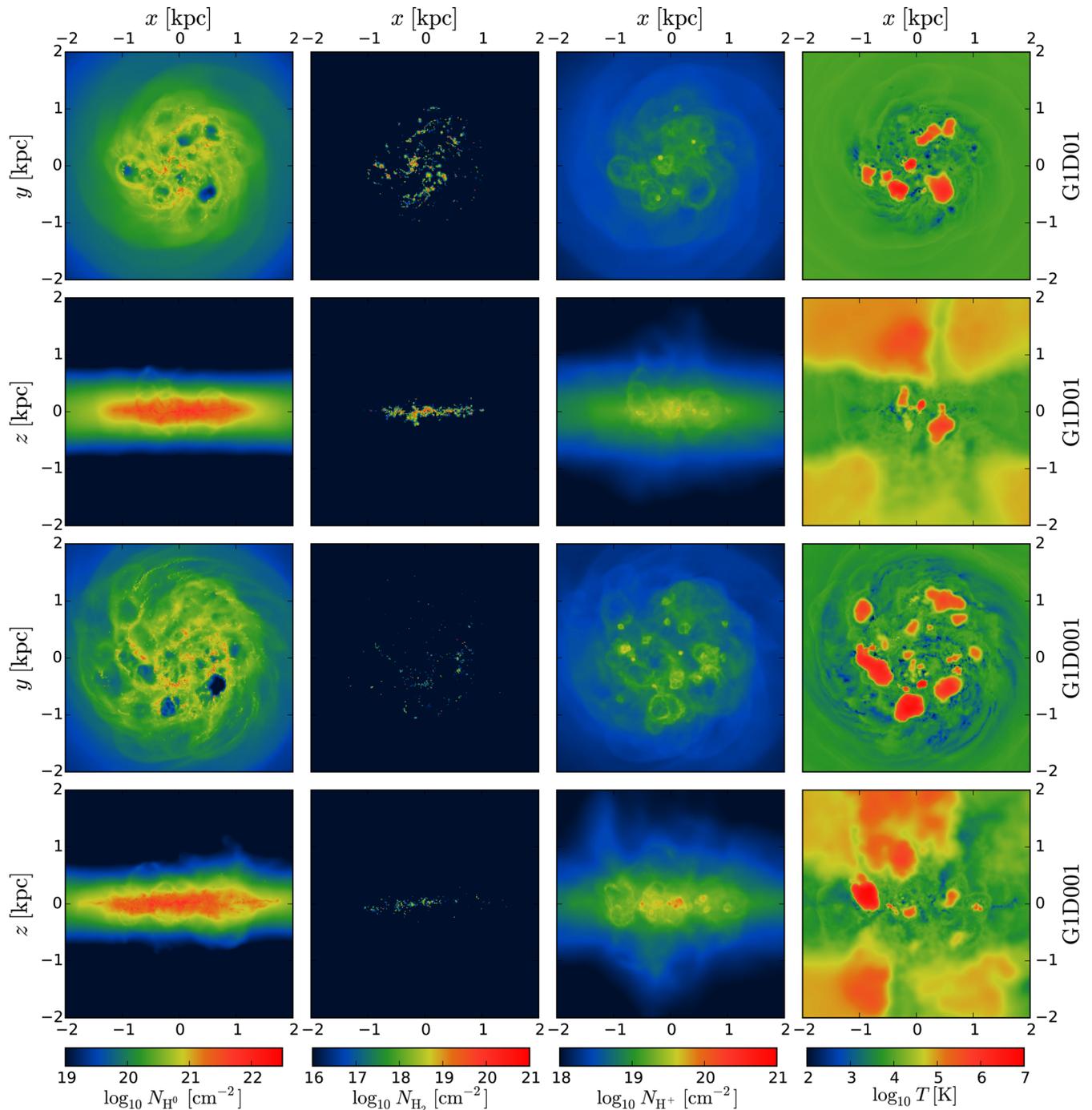
From the face-on images, the region where the density structures can be found is the largest in *G1D001* because the radius  $R$  beyond which gas cannot cool and collapse is the largest. This will be shown more quantitatively in Fig. 13.

In Fig. 5, we show the face-on column density for the *G1D001* run at  $t = 500 \text{ Myr}$  with zoom-in’s at different scales. The central large panel shows the disc within  $R < 2 \text{ kpc}$  and the other four panels are zoom-in views of the central one. The top-left panel shows a filamentary structure about 300 pc long. The bottom-left panel is an example of an SN bubble with a size of about 200 pc. The top-right panel shows a region with plenty of dense clouds and the bottom-right panel is a further zoom-in from the top right. Note that the spatial resolution we have in dense gas is about 2 pc and therefore the clumps in the bottom-right panel are expected to be still well resolved. The ISM is highly inhomogeneous with a complex density structure.

### 4.3 Time evolution of global properties

#### 4.3.1 SFR and $H_2$ mass fraction

In the upper panel of Fig. 6, we show the time evolution of the total SFR of the galaxy (SFR) for *D1G01*, *D1G001*, *D01G01* and *D1G01\_noFB*. In the first 50 Myr, the gas collapses on to a thin disc

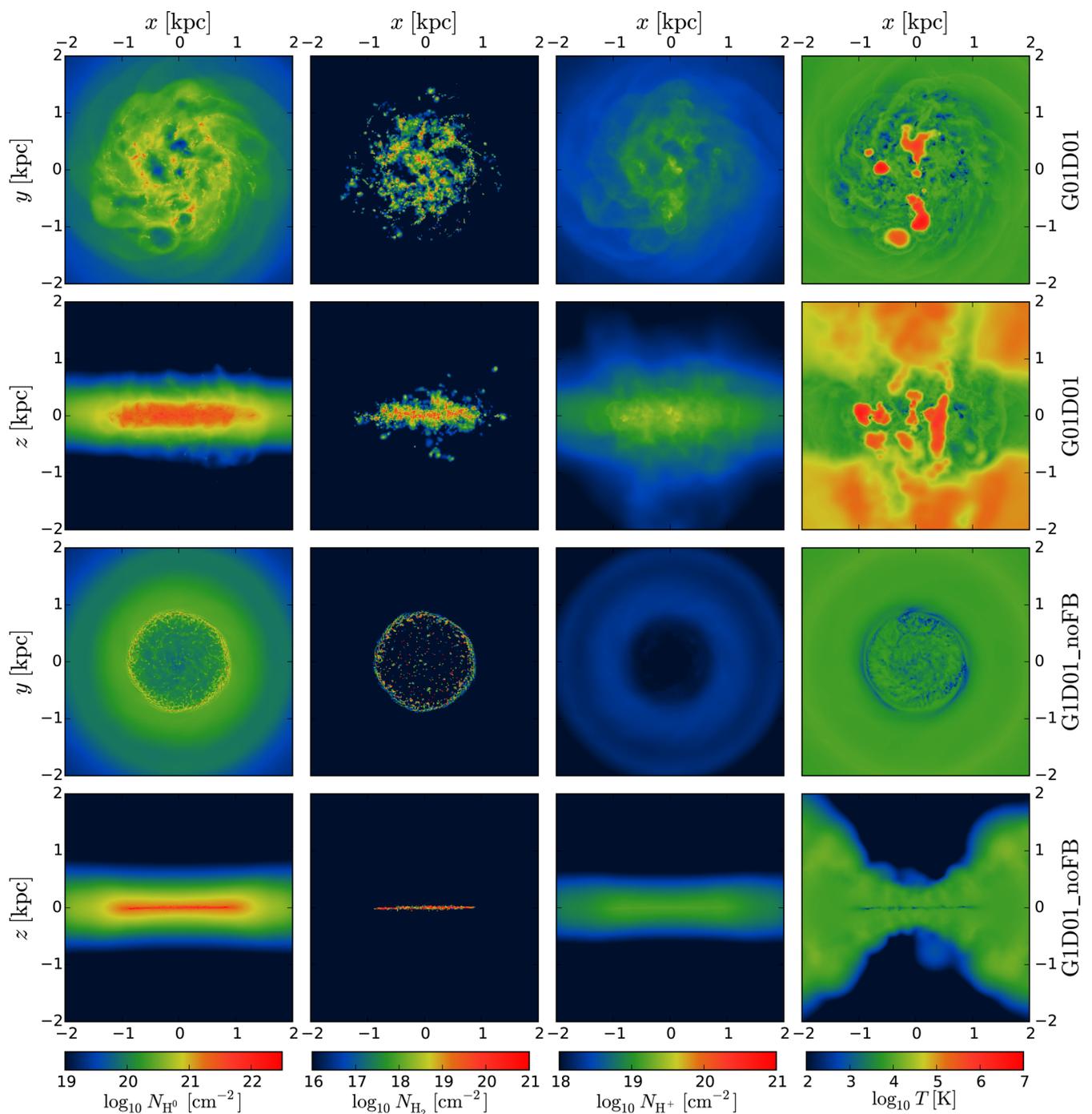


**Figure 3.** Images of *G1D01* and *G1D001* at  $t = 500$  Myr. From left to right: the column density maps of H I (first column), H<sub>2</sub> (second column) and H<sup>+</sup> (third column) and temperature maps (slices) (fourth column). The top two rows are the face-on and edge-on views for run *G1D01*, while the bottom two rows are for run *G1D001*. Note that we use different ranges of the colour bars for different chemical species for display purpose, as the system is strongly dominated by H I.

as it cools down and starts forming stars. This phase of global collapse is quickly terminated by SN feedback and the system gradually settles to a steady state, with  $\text{SFR} \approx 10^{-3} M_{\odot} \text{ yr}^{-1}$ . The gas depletion time of the galaxy  $t_{\text{dep}} \equiv M_{\text{gas}}/\text{SFR} \approx 40$  Gyr is so long that the gas reservoir is able to sustain the SFR throughout the simulation time. This is true even if galactic outflows are considered, as we will show in Section 4.3.3. The evolution of the SFRs is very similar in all runs with feedback, suggesting that the thermal properties of gas are not very sensitive to the strength of the ISRF

or the DGR. The only run that shows significant difference is the no-feedback one *G1D1\_noFB*. Here the SFR first gets as high as  $0.07 M_{\odot} \text{ yr}^{-1}$  at  $t = 200$  Myr, more than an order of magnitude higher than all the other runs, and quickly declines afterwards. This is due to a relatively short gas depletion time  $t_{\text{dep}} \approx 500$  Myr. The gas reservoir is rapidly consumed by star formation.

In the lower panel of Fig. 6 we show the H<sub>2</sub> mass fraction of the galaxy  $F_{\text{H}_2}$  (solid lines), defined as the total H<sub>2</sub> mass divided by the total gas mass in the ISM ( $F_{\text{H}_2} = \Sigma_i (f_{\text{H}_2}^i m_{\text{gas}}^i) / \Sigma_i m_{\text{gas}}^i$ ,

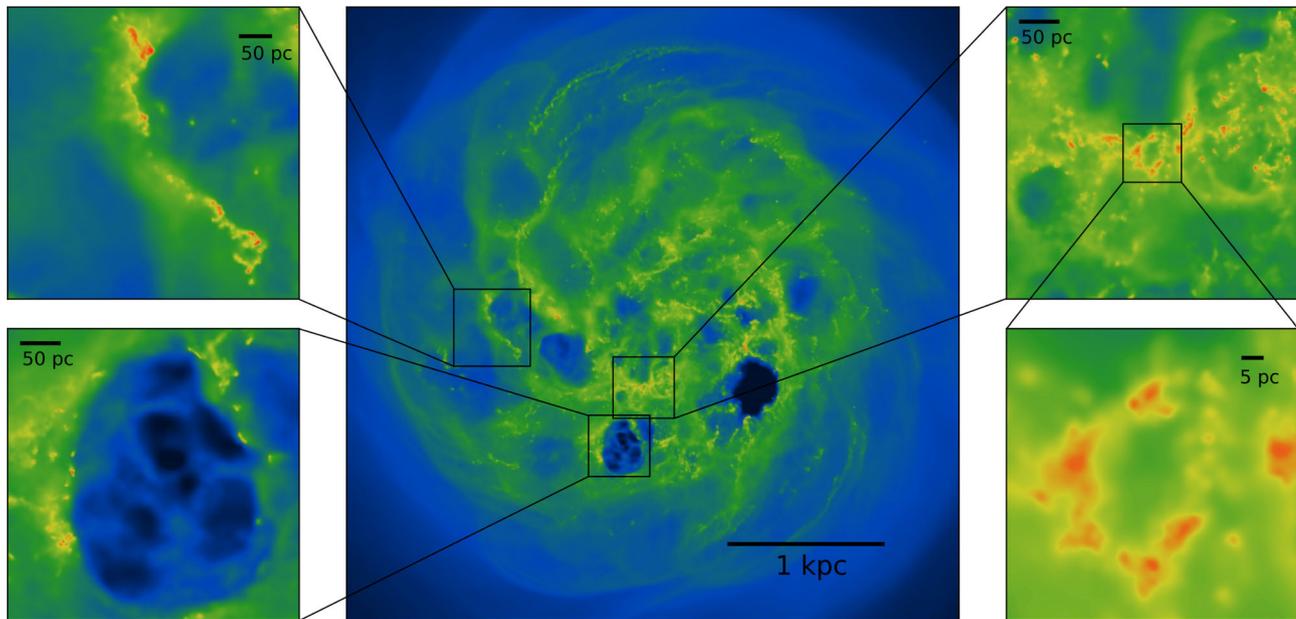


**Figure 4.** Same as Fig. 3 but for run *G01D01* (upper two rows) and run *G1D01\_noFB* (lower two rows).

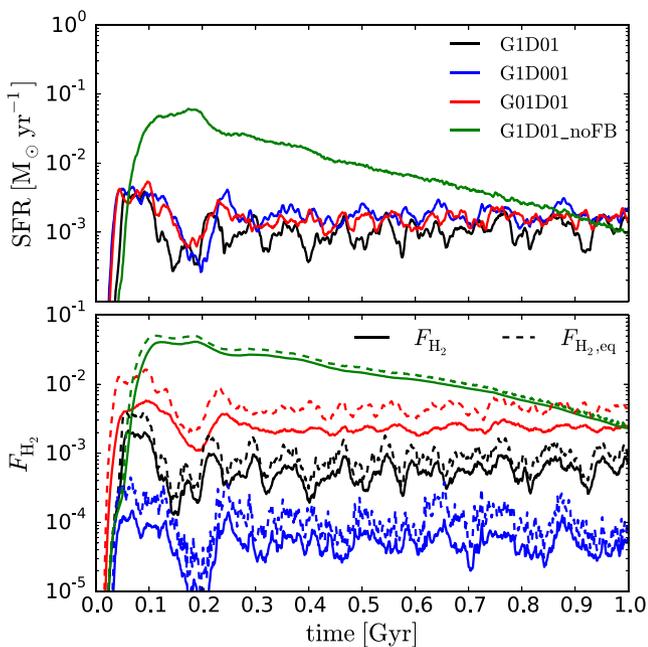
where  $m_{\text{gas}}^i$  and  $f_{\text{H}_2}^i$  are the mass and the  $\text{H}_2$  mass fraction of individual particle  $i$ ). The ISM is defined as gas within the region  $R < 2$  kpc (which is roughly the edge of star formation activities, see Fig. 13) and  $|z| < 1$  kpc (to exclude the halo region). The dashed lines show the corresponding (chemical-)equilibrium  $\text{H}_2$  mass fraction  $F_{\text{H}_2, \text{eq}} = \Sigma_i (f_{\text{H}_2, \text{eq}}^i m_{\text{gas}}^i) / \Sigma_i m_{\text{gas}}^i$ , where  $f_{\text{H}_2, \text{eq}}^i$  is calculated by equation (13) (i.e. assuming all particles are in chemical equilibrium). Note that the self-shielding factor is still obtained from the non-equilibrium  $f_{\text{H}_2}$  from the simulations. Unlike the SFR,  $F_{\text{H}_2}$  differs significantly between different runs. In *G1D01*  $F_{\text{H}_2}$  is about 0.05 per cent. Lowering  $D$  by a factor of 10 (*G1D001*) decreases

$F_{\text{H}_2}$  by more than an order of magnitude, as the  $\text{H}_2$  formation rate is directly proportional to  $D$ . With 10 times smaller  $G_0$  (*G01D01*) the difference is slightly smaller:  $F_{\text{H}_2}$  increases by about a factor of 3. The no-feedback run (*G1D1\_noFB*) shows the highest  $F_{\text{H}_2}$  (up to 4 per cent) as the global gravitational collapse forms massive dense clumps which provides an ideal environment for forming  $\text{H}_2$ . After  $t = 200$  Myr,  $F_{\text{H}_2}$  declines because most of the massive clumps have turned into stars.

Comparing  $F_{\text{H}_2}$  and  $F_{\text{H}_2, \text{eq}}$  in each run indicates that the equilibrium prediction systematically overestimates the  $\text{H}_2$  mass fraction. This difference is expected due to the slow  $\text{H}_2$  formation rate:



**Figure 5.** The face-on column density maps of H I for the *G1D001* run at  $t = 500$  Myr at different scales. The colour scale is the same as in Figs 3 and 4. Central panel: the entire star-forming region (box size = 4 kpc). Top left: a filamentary structure that is about 300 pc long. Bottom left: a 200-pc scale bubble driven by SN feedback. Top right: a group of dense clouds. Bottom right: further zoom-in of the top right. The effective spatial resolution is about 2 pc (see Section 2.8). The ISM is highly inhomogeneous with complex density structures.



**Figure 6.** Time evolution of global quantities of the galaxy. Upper panel: total SFR. The SFR is relatively insensitive to the choice of  $G_0$  and  $D$  but differs strongly if stellar feedback is switched off. Lower panel:  $H_2$  mass fraction in the ISM. The ISM is defined as gas within the region  $R < 2$  kpc and  $|z| < 1$  kpc. Solid lines ( $F_{H_2}$ ) are from the simulations while dashed lines ( $F_{H_2,eq}$ ) are calculated assuming chemical equilibrium. The  $H_2$  mass fraction is very sensitive to variations of  $G_0$  and  $D$ . The equilibrium prediction  $F_{H_2,eq}$  systematically overestimates the  $H_2$  mass fraction.

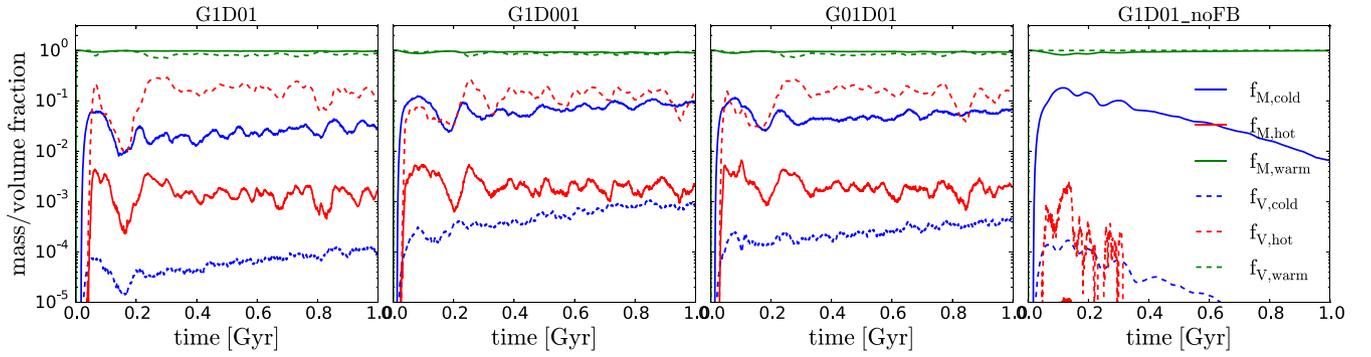
$t_F = (n_1 D)^{-1}$  Gyr. In the ISM constantly stirred by the SN-driven turbulence, it is unlikely for a gas cloud to stay unperturbed for a few  $t_F$  and form  $H_2$ . On the other hand, the destruction rate can also be slow in well-shielded regions, as shown in Section 3.1.1, which

also leads to some overabundant  $H_2$  gas that should have been destroyed if it were in chemical equilibrium. The overabundant and underabundant  $H_2$  gas compensate each other and therefore the global  $H_2$  fraction  $F_{H_2}$  is only slightly lower than  $F_{H_2,eq}$ , although locally they are far out of chemical equilibrium (as we will show in Section 4.4.2).

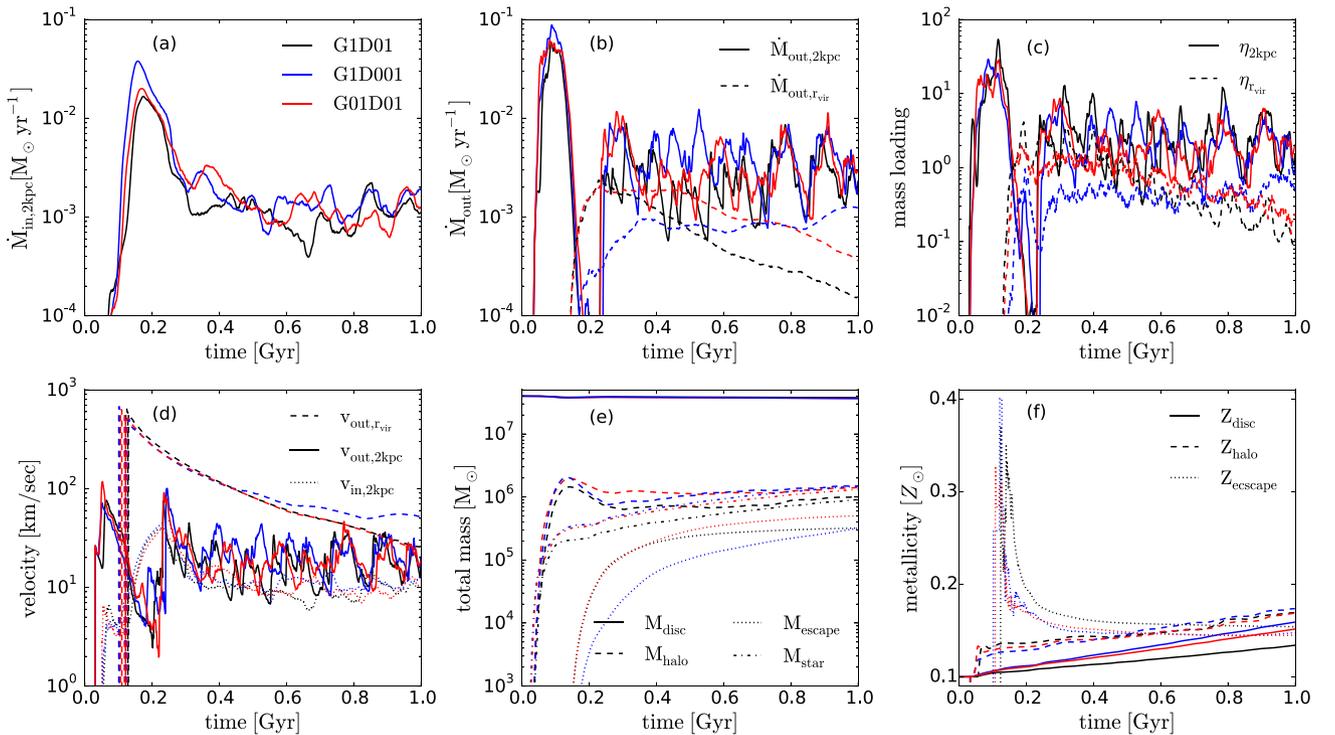
In the no-feedback run (*G1D01\_noFB*)  $F_{H_2}$  and  $F_{H_2,eq}$  agree pretty well. This is partly due to the relatively ‘quiescent’ ISM (due to the lack of SN feedback) allowing for  $H_2$  formation in a less disturbed environment. More importantly, the high-density clumps result in much shorter  $t_F$ , driving the system towards chemical equilibrium much faster.

#### 4.3.2 Mass and volume fraction of different phases

In Fig. 7, we show the time evolution of the mass fraction ( $f_M$ , solid lines) and volume fraction ( $f_V$ , dashed lines) of the ISM ( $R < 2$  kpc and  $|z| < 1$  kpc) in hot ( $T > 3 \times 10^4$  K), warm ( $100 \text{ K} < T \leq 3 \times 10^4$  K) and cold ( $T \leq 100$  K) phases. The volume fraction is obtained by integrating the volume (estimated by the smoothing kernel size,  $H^3$ )-weighted temperature histogram for different phases. The warm gas dominates the ISM both in mass ( $f_{M,warm} \approx 1$ ) and also in volume ( $f_{V,warm} \approx 0.9$ ). As visualized in Figs 3 and 4, the hot phase occupies a non-negligible volume fraction ( $f_{V,hot} \approx 0.1$ ), though it contributes only  $\approx 10^{-3}$  in mass. On the other hand, the volume occupied by the cold gas is negligibly small, though its mass fraction can be as high as  $f_{M,cold} \approx 0.1$ . The cold-gas fraction is a bit higher in *G1D001* and *G01D01* due to less efficient photoelectric heating in the low-DGR and weak-ISRF conditions (equation 6). In principle, a lower DGR also results in less dust-shielding, which would potentially result in less cold gas. However, dust-shielding only operates at  $N_{H,tot} > 1/(D\sigma_{dust})$ , which is rare in our simulations except for the densest gas. There is also a slight increase in the cold-gas fraction over time as a result of metal enrichment, which will be shown in Fig. 8. In the no-feedback run *G1D01\_noFB*, there is almost no hot gas as the hot



**Figure 7.** Time evolution of the mass fraction ( $f_M$ , solid lines) and volume fraction ( $f_V$ , dashed lines) of the ISM ( $R < 2$  kpc and  $|z| < 1$  kpc) in hot ( $T > 3 \times 10^4$  K), warm ( $100 \text{ K} < T \leq 3 \times 10^4$  K) and cold ( $T \leq 100$  K) phases. The warm gas dominates the ISM both in mass and in volume. In all runs except for *G1D01\_noFB*, the hot gas occupies a non-negligible volume ( $f_{V,\text{hot}} \approx 0.1$ ) but contains little mass, while the cold gas shows the opposite behaviour and slightly increases over time due to metal enrichment. In *G1D01\_noFB*, the hot gas is absent and the cold gas decreases over time due to the rapid depletion of gas by star formation.



**Figure 8.** Time evolution of outflow-related physical quantities in *G1D01* (black), *G1D001* (blue) and *G01D01* (red). Panel (a): the inflow rate at  $|z| = 2$  kpc ( $\dot{M}_{\text{in},2\text{kpc}}$ ). Panel (b): the outflow rate at  $|z| = 2$  kpc ( $\dot{M}_{\text{out},2\text{kpc}}$ ) and at  $r = r_{\text{vir}}$  ( $\dot{M}_{\text{out},r_{\text{vir}}}$ ). Panel (c): the mass-loading factor at  $|z| = 2$  kpc ( $\eta_{2\text{kpc}}$ ) and at  $r = r_{\text{vir}}$  ( $\eta_{r_{\text{vir}}}$ ). Panel (d): the mass-weighted outflow velocity at  $|z| = 2$  kpc ( $v_{\text{out},2\text{kpc}}$ ) and at  $r = r_{\text{vir}}$  ( $v_{\text{out},r_{\text{vir}}}$ ), and inflow velocity at  $|z| = 2$  kpc ( $v_{\text{in},2\text{kpc}}$ ). Panel (e): the total gas mass in the disc ( $M_{\text{disc}}$ ), in the halo ( $M_{\text{halo}}$ ), outside of the halo ( $M_{\text{escape}}$ ), and the total mass of stars formed in the simulations ( $M_{\text{star}}$ ). Panel (f): the mass-weighted metallicity in the disc ( $Z_{\text{disc}}$ ), in the halo ( $Z_{\text{halo}}$ ) and outside of the halo ( $Z_{\text{escape}}$ ).

gas is mostly produced by SN blastwaves, and the cold-gas fraction decreases over time due to the high SFR depleting the reservoir.

#### 4.3.3 Feedback-driven galactic outflows

SN explosions inject energy into the ISM and may push the gas out of the disc and drive galactic outflows. We define the outflow rate as the mass flux integrated over a chosen surface  $S$ :

$$\dot{M}_{\text{out}} \equiv \int_S \rho \mathbf{v} \cdot \hat{\mathbf{n}} da \quad (15)$$

where  $\rho$  is the gas density,  $\mathbf{v}$  is the gas velocity and  $\hat{\mathbf{n}}$  is the normal direction of the surface (in the obvious sense of outwards from the disc). In SPH simulations,  $\dot{M}_{\text{out}}$  can be estimated as  $\dot{M}_{\text{out}} = \sum_i m_{\text{gas}}^i \mathbf{v}^i \cdot \hat{\mathbf{n}} / \Delta x$  where  $m_{\text{gas}}^i$  and  $\mathbf{v}^i$  are the mass and velocity of particle  $i$ , and  $\Delta x$  is a pre-defined thickness of the measuring surface. The summation is over particles within the shell of the surface with  $\mathbf{v}^i \cdot \hat{\mathbf{n}} > 0$ . The choice of  $\Delta x$  is a compromise between being too noisy ( $\Delta x$  too small) and oversmoothing ( $\Delta x$  too large). The inflow rate  $\dot{M}_{\text{in}}$  is defined in the same way but with the direction of  $\hat{\mathbf{n}}$  reversed.

As a proxy for how much gas is leaving or entering the disc, we measure the outflow rate  $\dot{M}_{\text{out},2\text{kpc}}$  and inflow rate  $\dot{M}_{\text{in},2\text{kpc}}$  at the

planes  $z = \pm 2$  kpc, truncated at  $R < 6$  kpc. The choice of  $z = \pm 2$  kpc is to avoid counting the thick disc material as outflows/inflows, especially at larger radii where the gas disc flares, (though the distinction between disc material and outflows/inflows is somewhat arbitrary). The thickness of the planes  $\Delta x$  is set to 0.2 kpc but the results are not sensitive to the chosen value. To assess how much gas is escaping the dark matter halo, we also measure the outflow/inflow rate at the virial radius  $r = 44$  kpc (denoted as  $\dot{M}_{\text{out/in,vir}}$ ). Here the thickness of the sphere  $\Delta x$  is set to 2 kpc due to the much lower density around the virial radius. The mass-loading factor  $\eta$  is defined as  $\dot{M}_{\text{out}}/\text{SFR}$ . There should be a time difference between when the star formation occurs and when the outflowing material causally linked to that star formation event reaches the measuring surface. Our definition of the mass-loading factor does not account for this effect and thus may seem inappropriate. This is especially true for  $\eta_{\text{vir}}$  as it takes about 200 Myr for the gas to reach  $r_{\text{vir}}$  assuming the outflowing velocity of  $200 \text{ km s}^{-1}$ . However, as the SFR does not change dramatically throughout the simulations,  $\eta$  can still be a useful proxy for how efficient feedback is at driving outflows.

In Fig. 8 we show the time evolution of  $\dot{M}_{\text{in,2kpc}}$  in panel (a),  $\dot{M}_{\text{out,2kpc}}$  and  $\dot{M}_{\text{out,vir}}$  in panel (b), and  $\eta_{2\text{kpc}}$  and  $\eta_{\text{vir}}$  in panel (c), respectively. First, we focus on the  $z = \pm 2$  kpc planes. Comparing with Fig. 6,  $\dot{M}_{\text{out,2kpc}}$  rises sharply right after the first star formation occurs. Since there is no gas above the disc in the initial conditions, the initial outflows gush out freely into vacuum and thus the mass loading is very high, ( $\eta_{2\text{kpc}} > 10$ ). This initial phase of violent outflow lasts about 100 Myr. A significant fraction of gas falls back on to the disc due to the gravitational pull, which gives rise to the inflows starting from  $t \approx 150$  Myr. From  $t \gtrsim 200$  Myr, the system reaches a quasi-steady state with a small but non-zero net outflow  $\dot{M}_{\text{out,2kpc}} - \dot{M}_{\text{in,2kpc}}$ . The large fluctuations of  $\dot{M}_{\text{out,2kpc}}$  indicate that the outflows are intermittent on time-scales of about 50 Myr. On the time-scale of 1 Gyr, however,  $\dot{M}_{\text{out,2kpc}}$  is rather constant and is slightly higher than the SFR ( $\eta_{\text{out,2kpc}} \gtrsim 1$ ).

Now we turn to the spherical surface at  $r_{\text{vir}} = 44$  kpc. Due to the time delay for gas to travel from the disc to  $r_{\text{vir}}$ , the first non-zero  $\dot{M}_{\text{out,vir}}$  appears at  $t \approx 200$  Myr. The peak value of  $\dot{M}_{\text{out,vir}}$  is much lower than the peak of  $\dot{M}_{\text{out,2kpc}}$  at  $t \approx 80$  Myr. While this is partly due to the gravitational pull that slows down the outflowing gas, the more important reason is that not all of the gas which passed through the  $z = \pm 2$  kpc planes eventually made it to  $r_{\text{vir}}$ . Some would later fall back on to the disc (galactic fountain) and some would fill the space in the halo region. No inflows at  $r_{\text{vir}}$  are detected ( $\dot{M}_{\text{in,vir}} = 0$  at all times) so  $\dot{M}_{\text{out,vir}}$  can be regarded as a net outflow. The mass-loading  $\eta_{\text{out,vir}}$  first reaches about the order of unity and then decreases as the gas accumulates in the halo and hinders the subsequent outflows.

In Fig. 8 panel (d), we show the averaged velocity  $v_{\text{out,vir}}$  at  $r_{\text{vir}}$  and  $v_{\text{out/in,2kpc}}$  at  $z = \pm 2$  kpc, respectively. The averaged velocity is mass-weighted over the particles within the measuring shells. It may seem inconsistent that  $v_{\text{out,vir}}$  is much larger than  $v_{\text{out,2kpc}}$ . This is not because the outflowing particles are accelerated, but is again due to the fact that not all particles crossing the  $z \pm 2$  kpc planes have enough kinetic energy to eventually escape the halo, and the averaged velocities are thus much smaller. The initial high  $v_{\text{out,vir}}$  is due to the vacuum initial conditions in the halo region, and over time  $v_{\text{out,vir}}$  decreases as the halo gas hinders the subsequent outflows.

In Fig. 8 panel (e), we show the total gas mass in different spatial regions as a function of time. The disc region is defined as  $R < 6$  kpc and  $|z| < 2$  kpc, and the halo is defined as  $r < r_{\text{vir}}$  excluding the disc region. We denote  $M_{\text{disc}}$ ,  $M_{\text{halo}}$  and  $M_{\text{escape}}$  as the total gas mass in the disc, in the halo, and outside of the halo, respectively. The

total mass of stars that have formed in the simulations is denoted as  $M_{\text{star}}$  and is well confined within the disc region. Except for the initial phase of strong outflow, there is roughly the same amount of gas ejected into the halo as the amount of stars formed in the simulations ( $M_{\text{halo}} \approx M_{\text{star}}$ ), both of which are more than one order of magnitude lower than  $M_{\text{disc}}$ . Therefore,  $M_{\text{disc}}$  hardly changes throughout the simulations, and as a consequence the system is able to settle into a quasi-steady state. The escaped mass  $M_{\text{escape}}$  is about 10–30 per cent of  $M_{\text{halo}}$  (and  $M_{\text{star}}$ ), which is consistent with  $\eta_{\text{vir}} < 1$  shown in panel (c).

In Fig. 8 panel (f), we show the mass weighted mean metallicity of the gas in the disc,  $Z_{\text{disc}}$ , in the halo,  $Z_{\text{halo}}$ , and for gas that has escaped the halo,  $Z_{\text{escape}}$ . Starting from  $0.1 Z_{\odot}$ , the disc metallicity slowly increases due to metal enrichment from the stellar feedback. At the end of simulations ( $t = 1$  Gyr),  $Z_{\text{disc}}$  only increased to about  $0.15 Z_{\odot}$  in *G1D01* and *G01D01* and about  $0.12 Z_{\odot}$  in *G1D01*. The slightly slower enrichment in *G1D01* is due to its lower SFR, a consequence of more efficient photoelectric heating that suppresses gas cooling and star formation. The metallicity in the halo  $Z_{\text{halo}}$  is only slightly higher (about 20 per cent) than  $Z_{\text{disc}}$ , as not only the highly enriched SN-ejecta but also the low-metallicity gas in the ISM is pushed out of the disc, and  $Z_{\text{halo}}$  is thus diluted by the low-metallicity gas. The metallicity of the escaped gas  $Z_{\text{escape}}$  shows a sharp peak as high as  $0.3 Z_{\odot}$  initially, but also drops to about the same level as  $Z_{\text{halo}}$  very quickly (less than 50 Myr). In general, the outflows in our simulations are enriched winds, moderately.

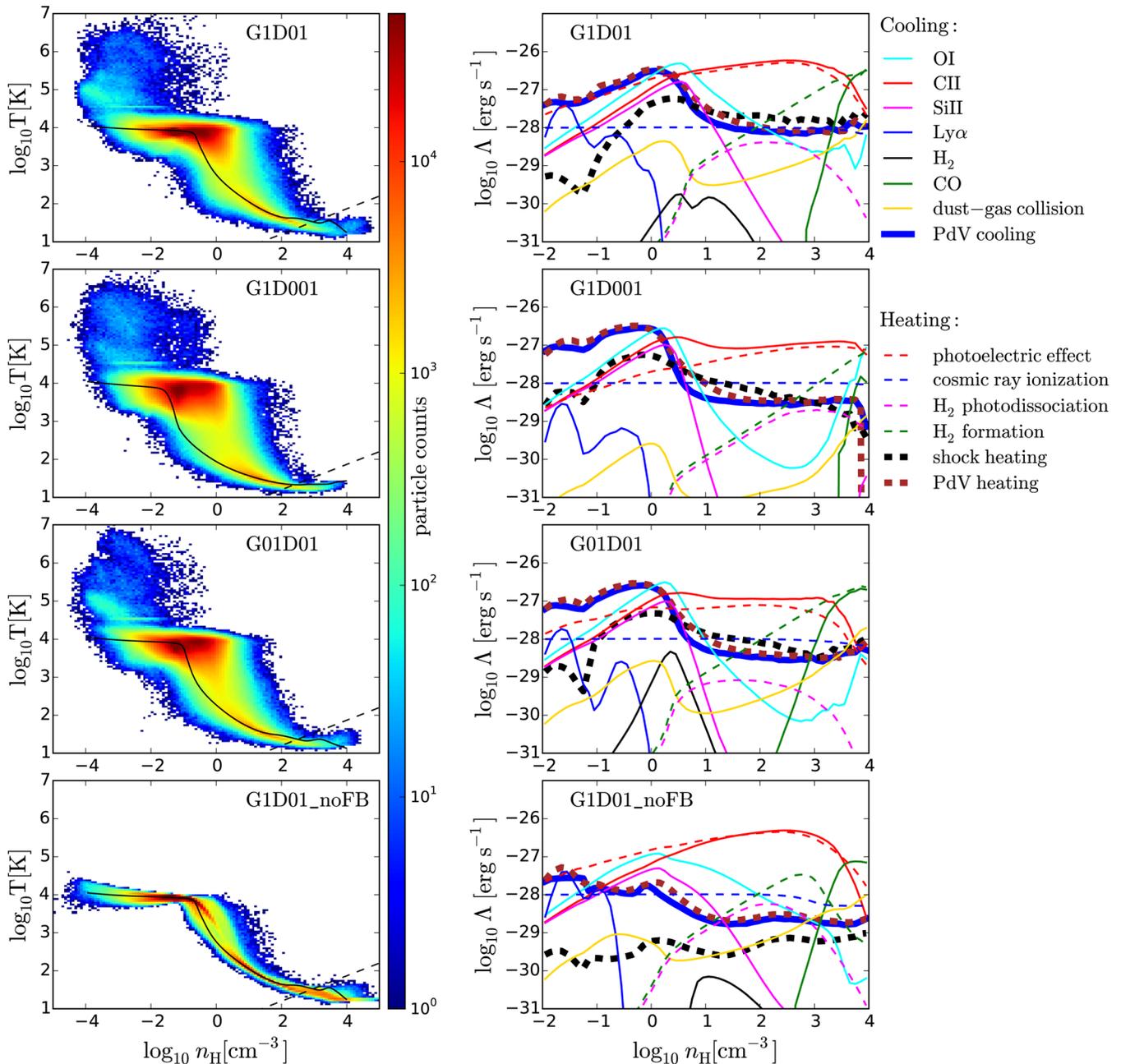
#### 4.4 ISM properties

Having described the global properties of the simulated galaxies, we now turn to the local properties of the ISM and its multiphase structure.

##### 4.4.1 Thermal properties of the ISM

In Fig. 9, we show the phase diagrams (temperature versus density, left-hand panels) and the corresponding cooling and heating rates of different processes (median value, right-hand panels) in the ISM region ( $R < 2$  kpc and  $|z| < 1$  kpc) at  $t = 300$  Myr in run *G1D01*, *G1D001*, *G01D01* and *G1D01\_noFB* (from top to bottom). The black curves in the left-hand panels show the (thermal-)equilibrium temperature–density relation as shown in Fig. 2. The dashed lines show the contours where  $M_J = M_{\text{ker}}$ . Particles lying below the dashed lines are unresolved in terms of self-gravity. Except for *G1D1\_noFB* where the gas has undergone runaway collapse, the majority of gas lies in the Jeans-resolved region up to  $n_{\text{H}} \approx 200 \text{ cm}^{-3}$ . In the right-hand panels, the thick dashed black curves represent the shock-heating rate, i.e. the viscous heating. Note that this shock-heating rate does not include the thermal injection of the SN feedback, which instantaneously heats up the nearest 100 gas particles as described in Section 2.6.1. The thick solid blue and dashed brown lines show, respectively, the cooling and heating rate caused by adiabatic expansion and compression  $\dot{u}_{\text{p,dv}} = -(2/3)u \nabla \cdot \mathbf{v}$  where  $u$  is the specific internal energy and  $\nabla \cdot \mathbf{v}$  is the velocity divergence.

For  $n_{\text{H}} \geq 100 \text{ cm}^{-3}$ , the gas is approximately in thermal equilibrium at around few tens of Kelvin as the cooling time is much shorter than its local dynamical time (can be estimated by  $\rho/\dot{\rho} = (3/2)u/\dot{u}_{\text{p,dv}} = -(\nabla \cdot \mathbf{v})^{-1}$ ). A small bump at the highest densities is due to  $\text{H}_2$  formation heating. Below  $n_{\text{H}} = 100 \text{ cm}^{-3}$ , the scatter starts to increase due to feedback-driven turbulent stirring. For  $n_{\text{H}} < 1 \text{ cm}^{-3}$ , the adiabatic heating/cooling dominates

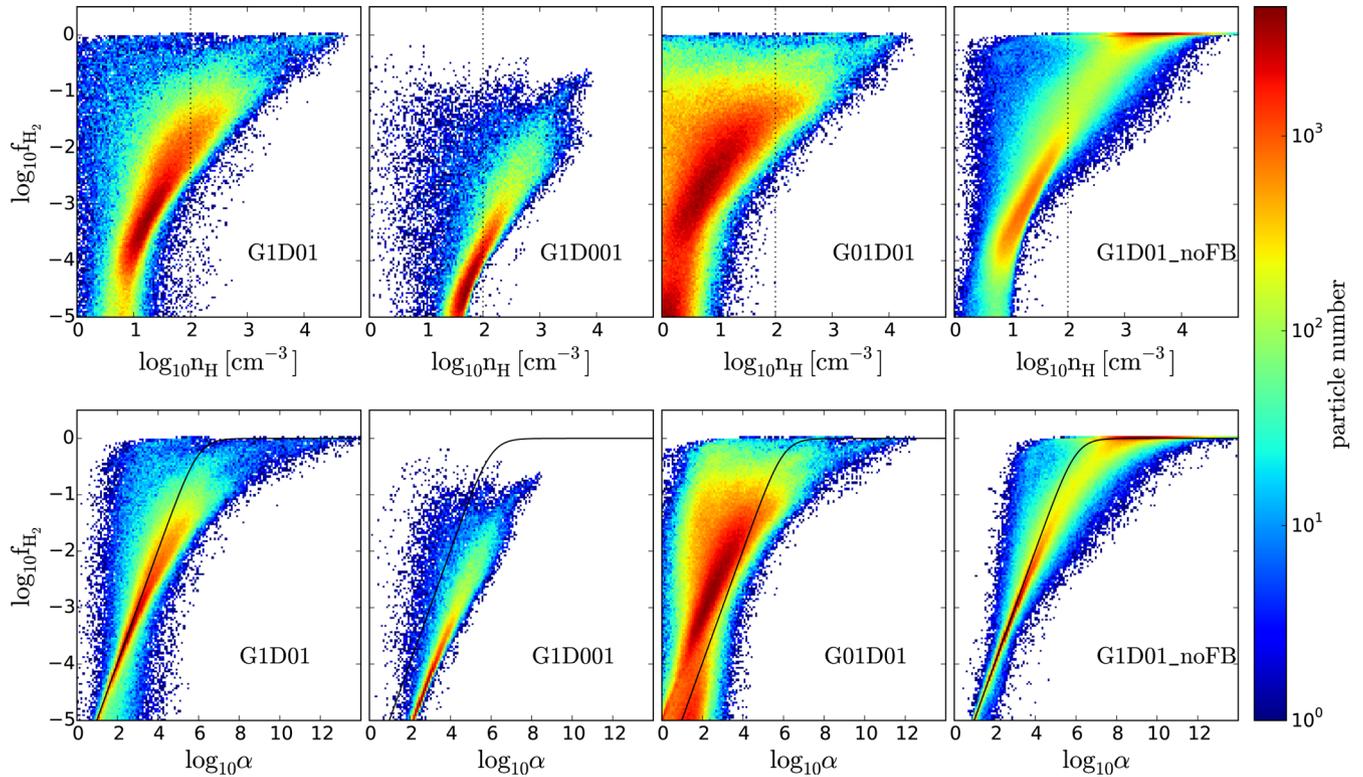


**Figure 9.** Phase diagrams (temperature versus density, left-hand panels) and the cooling and heating rates of different processes (median value, right-hand panels) in the ISM region ( $R < 2$  kpc and  $|z| < 1$  kpc) at  $t = 300$  Myr in run *G1D01*, *G1D001*, *G01D01* and *G1D01\_noFB*. The black solid lines shows the (thermal)-equilibrium temperature–density relation as shown in Fig. 2. The dashed line shows the contour where  $M_J = M_{\text{ker}}$ . SN feedback keeps the gas temperatures at  $10^4$  K in the range of  $0.1 < n_H < 10 \text{ cm}^{-3}$  even if the equilibrium temperatures are much lower. As a result, the distribution is not sensitive to  $G_0$  and  $D$ , both of which affect photoelectric heating but not SN feedback.

over other radiative processes, and the gas spends several dynamical times before it can cool. Once the gas cools below  $10^4$  K, it is likely to be shock heated to  $10^4$  K in a dynamical time (gas being shock heated to  $T > 10^4$  K is possible but at such temperatures the gas will quickly cool down to  $10^4$  K). Therefore, turbulence drives the gas temperatures away from the equilibrium curve and keeps them at  $10^4$  K at  $n_H \lesssim 1 \text{ cm}^{-3}$  (see also Walch et al. 2011 and Glover & Clark 2012c).

The distribution of gas in the phase diagram is not very sensitive to the choice of  $G_0$  and  $D$ . This is because SN feedback, which

enhances turbulence and drives the gas out of thermal equilibrium in the range of  $0.1 < n_H < 10 \text{ cm}^{-3}$  where the equilibrium curve is most sensitive to  $G_0$  and  $D$  (Fig. 2), does not depend explicitly on  $G_0$  and  $D$ . In *G1D01*, the gas distribution follows the equilibrium curve better as the photoelectric heating is more efficient and dominates over shock heating. In *G1D1\_noFB*, the gas also follows the equilibrium curve and shows the least scatter, as the turbulent stirring is much weaker without SN feedback. There is no hot gas in *G1D1\_noFB* because the hot gas is only generated by the SN feedback.



**Figure 10.** Top panels: the  $\text{H}_2$  mass fraction  $f_{\text{H}_2}$  versus the number density  $n_{\text{H}}$  at  $t = 300$  Myr. The dotted line indicates the star formation threshold density  $n_{\text{H, th}} = 100 \text{ cm}^{-3}$ . Most star formation actually occurs in regions dominated by atomic hydrogen, while  $\text{H}_2$ -dominated regions are not necessarily star forming. Bottom panels:  $f_{\text{H}_2}$  versus the dimensionless quantity  $\alpha \equiv n_1 D / f_{\text{sh}} G_0$  at  $t = 300$  Myr. The solid line shows the  $f_{\text{H}_2}$ - $\alpha$  relation in chemical equilibrium (cf. equation 14). Particles with under/overabundant  $f_{\text{H}_2}$  lie below/above the black curves. The gas is far from chemical equilibrium locally.

#### 4.4.2 Molecular hydrogen

In the top row of Fig. 10, we show the  $\text{H}_2$  mass fraction  $f_{\text{H}_2}$  versus the number density  $n_{\text{H}}$ . The dashed lines represent the star formation threshold  $n_{\text{H, th}} = 100 \text{ cm}^{-3}$ . Since most gas particles with  $n_{\text{H}} > n_{\text{th}}$  are also cooler than  $100 \text{ K}$  (cf. Fig. 9), we can regard those located to the right of dashed lines as star-forming gas, while those to the left are not star forming. In general,  $f_{\text{H}_2}$  increases with  $n_{\text{H}}$  as higher density leads to a higher  $\text{H}_2$  formation rate and usually implies more shielding. Except for *G1D01\_noFB*, a large fraction of star-forming gas has low  $f_{\text{H}_2}$ . The  $\text{H}_2$  mass fraction of gas with  $n_{\text{H}} > 100 \text{ cm}^{-3}$  is  $F_{\text{H}_2, \text{sfg}} = 10, 0.7, 18$  and  $70$  per cent in *G1D01*, *G1D001*, *G01D01* and *G1D01\_noFB*, respectively, though the precise values depend on our definition of star-forming gas ( $n_{\text{H, th}}$ ). This implies that most star formation actually occurs in regions dominated by atomic hydrogen (except for *G1D01\_noFB*), in agreement with previous theoretical predictions (Glover & Clark 2012c).

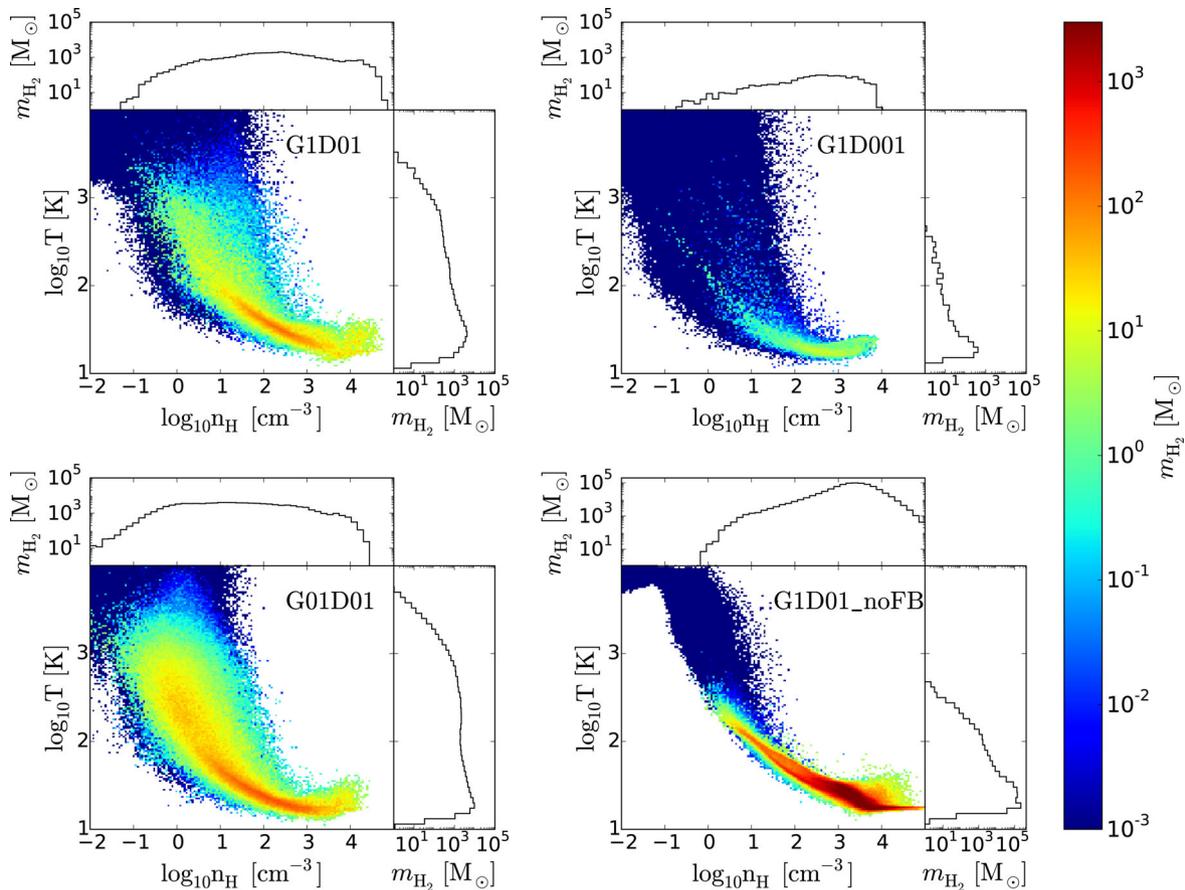
On the other hand, there are also particles whose densities are too low to be star forming but that have  $f_{\text{H}_2} \approx 1$ . The fraction of the total  $\text{H}_2$  mass that resides in  $n_{\text{H}} < 100 \text{ cm}^{-3}$  is  $42, 25, 70$  and  $4$  per cent in *G1D01*, *G1D001*, *G01D01* and *G1D01\_noFB*, respectively. This demonstrates that, just as in our own Galaxy,  $\text{H}_2$ -dominated regions are not necessarily star forming.<sup>3</sup> This is in line with Elmegreen &

Hunter (2015) who also concluded that there should be a significant amount of diffuse  $\text{H}_2$  in dwarf galaxies.

As discussed in Section 3.1.1, the formation and destruction time-scales of  $\text{H}_2$  are long compared to the free-fall time. In the presence of feedback, turbulence further reduces the local dynamical time. Therefore, the gas can easily be out of chemical equilibrium. In the bottom row of Fig. 10, we show  $f_{\text{H}_2}$  versus the dimensionless quantity  $\alpha \equiv n_1 D / f_{\text{sh}} G_0$  (see Section 3.1 for the definition). The black curves indicate the  $f_{\text{H}_2}$ - $\alpha$  relation in chemical equilibrium (equation 14). Gas particles with under/overabundant  $f_{\text{H}_2}$  lie below/above the black curves. The gas would follow the equilibrium curves if the chemical time-scales were much shorter than the local dynamical time-scales. The large scatter of the distribution indicates the opposite: the gas is far from chemical equilibrium locally. However, since there are both significant amount of particles underabundant and overabundant in  $\text{H}_2$ , the total  $\text{H}_2$  mass of the galaxy is not very different from the equilibrium predictions (cf. Fig. 6).

As shown in Fig. 6, the global value of  $F_{\text{H}_2}$  is sensitive to the assumed  $G_0$  and  $D$ . This is also true locally: the particle distribution in Fig. 10 differs dramatically when adopting different  $G_0$  and  $D$ . This is in contrast to the particle distribution in the phase diagram (Fig. 9) which is rather insensitive to variations of  $G_0$  and  $D$ . The largest scatter is in *G01D01*, as the weak ISRF leads to a lower  $\text{H}_2$  destruction rate. A significant fraction of overabundant  $\text{H}_2$  can therefore survive for a longer time. In *G1D001*, the low DGR makes it difficult to have high  $f_{\text{H}_2}$  even for the densest gas. Low  $f_{\text{H}_2}$  means little self-shielding (dust-shielding is almost absent given its DGR) and thus shorter  $\text{H}_2$  destruction time. In addition, although its  $\text{H}_2$  formation time is the longest among all runs, it takes much

<sup>3</sup> We note that if we were able to follow the collapse to much smaller scales and much higher densities (e.g. pre-stellar cores), eventually all star formation will occur in  $\text{H}_2$ . Our definition of star-forming gas corresponds to the reservoir gas that correlates (both spatially and temporary) with star formation.



**Figure 11.** Phase diagrams weighted by the  $H_2$  mass ( $f_{H_2}m_{\text{gas}}$ ) at  $t = 300$  Myr. The projected histograms of density and temperature (also weighted by  $f_{H_2}m_{\text{gas}}$ ) are shown on the upper and right-hand sides of each panel, respectively. Most  $H_2$  is located in the cold ( $T < 100$  K) gas. A significant fraction of  $H_2$  is located in the warm and diffuse phase in *G1D01* and *G01D01* due to non-equilibrium effects.

less time to reach the low-equilibrium mass fraction  $f_{H_2,\text{eq}}$ . The scatter in *G1D001* is therefore the smallest. Despite the absence of feedback, the *G1D1\_noFB* run also shows some scatter as a result of its turbulent motions due to thermal-gravitational instabilities. Nevertheless, there is a large fraction of  $H_2$ -dominated ( $f_{H_2} \approx 1$ ) gas that follows the equilibrium curve quite well. This explains the excellent agreement between the global  $f_{H_2}$  and the equilibrium prediction in Fig. 6.

To see in which phase most of the  $H_2$  can be found, we show in Fig. 11 the phase diagram weighted by the  $H_2$  mass. The projected histograms of density and temperature (also weighted by  $f_{H_2}m_{\text{gas}}$ ) are shown on the upper and right-hand sides of each panel, respectively. Most of  $H_2$  can be found in the cold and dense phase of the ISM. In *G1D01* and *G01D01* there is some  $H_2$  in the warmer ( $T \lesssim 1000$  K) and less dense phase. This is the  $H_2$ -overabundant, high  $f_{H_2}$  gas shown in Fig. 10, which would not be predicted by the equilibrium approach. Note that the  $H_2$  distribution does not necessarily reside in the star-forming regimes ( $T < 100$  K and  $n_H > 100 \text{ cm}^{-3}$ ). Again, this implies that non-star-forming regions can also be  $H_2$ -dominated.

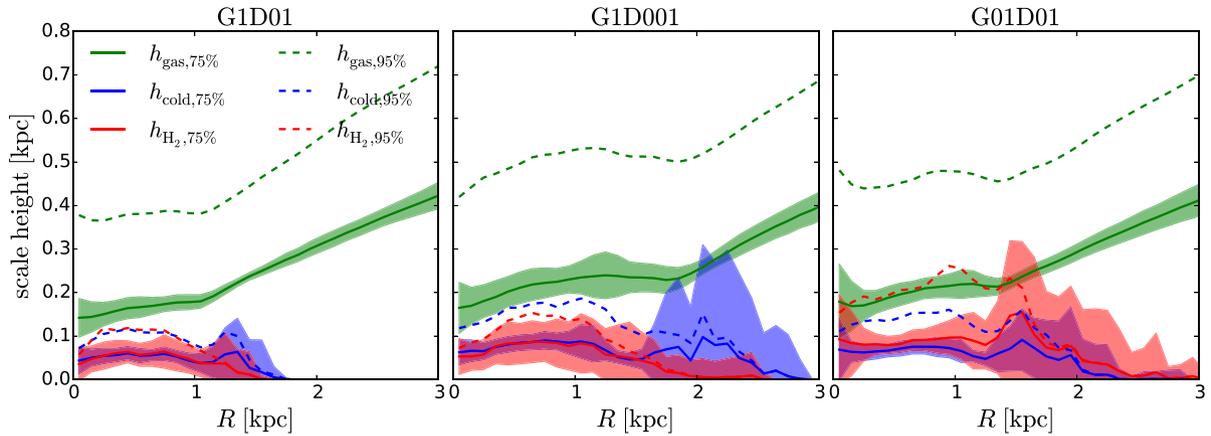
## 4.5 Radial variations

### 4.5.1 Scaleheight of the gas disc

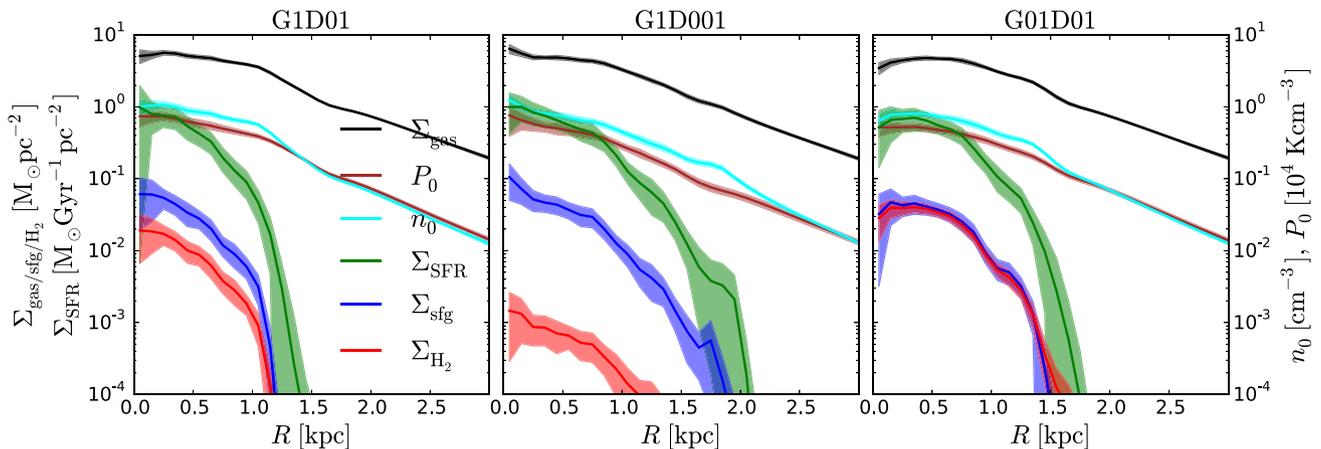
We define the scaleheight  $h_{\text{gas},75 \text{ per cent}}$  as the altitude where 75 per cent of the gas mass between  $z = \pm h_{\text{max}}$  is enclosed between  $z = \pm h_{\text{gas},75 \text{ per cent}}$  in each radial bin. The choice of  $h_{\text{max}} = 2$  kpc

is to discard the particles that clearly do not belong to the disc in any reasonable definition. Setting  $h_{\text{max}} = 1$  kpc gives essentially identical results. The definition of 75 per cent enclosed mass is motivated by the  $\text{sech}^2(z)$  distribution for a self-gravitating disc ( $\tanh(-1) \approx 0.75$ ). Similarly, we define  $h_{\text{gas},95 \text{ per cent}}$  in the same way as  $h_{\text{gas},75 \text{ per cent}}$  but for 95 per cent of enclosed mass (as  $\tanh(-2) \approx 0.95$ ). We also define the scaleheight of the cold ( $T < 100$  K) gas and the  $H_2$  in the same fashion, denoted as  $h_{\text{cold},75(95) \text{ per cent}}$  and  $h_{H_2,75(95) \text{ per cent}}$ , respectively.

In Fig. 12 we show  $h_{\text{gas},75/95 \text{ per cent}}$  (green),  $h_{\text{cold},75/95 \text{ per cent}}$  (blue) and  $h_{H_2,75/95 \text{ per cent}}$  (red) in the solid/dashed curves as a function of  $R$ , (arithmetic-)averaged over time from  $t = 250$  Myr to 1 Gyr with a time interval of 1 Myr. The filled areas show the  $\pm 1\sigma$  bands where  $\sigma$  is the standard deviation of  $h_{\text{gas}/\text{cold}/H_2,75 \text{ per cent}}$  in each radial bin. The size of each radial bin is 0.1 kpc. The scaleheight of the total gas  $h_{\text{gas},75 \text{ per cent}}$  and  $h_{\text{gas},95 \text{ per cent}}$  both increase with  $R$ , especially at larger radii. On the other hand, 75 per cent (95 per cent) of the cold gas and  $H_2$  is found within a thin layer  $|z| < 0.1(0.2)$  kpc without flaring, in line with what Elmegreen & Hunter (2015) inferred. There is some apparent flaring of  $h_{\text{cold}}$  and  $h_{H_2}$  for the  $1\sigma$  band (but not the mean value) at  $R \approx 2$  kpc in both *G1D01* and *G01D01*. However, very little cold gas and  $H_2$  can be found at such radii (cf. Fig. 13). Star formation mainly takes place in the mid-plane  $|z| < 0.1$  kpc, as this is where the gas density is high enough to cool efficiently and trigger gravitational collapse. The cold and dense environment in the mid-plane also favours  $H_2$  formation. Once the  $H_2$  is pushed above the disc by SN feedback, it would be destroyed by the



**Figure 12.** Scaleheight of the disc as a function of  $R$ .  $h_{\text{gas},75\text{ per cent}}$  (95 per cent) is defined as the altitude where 75 per cent (95 per cent) of the gas mass between  $z = \pm h_{\text{max}}$  is enclosed between  $z = \pm h_{\text{gas},75\text{ per cent}}$  (95 per cent). The scaleheight of the cold ( $T < 100$  K) gas  $h_{\text{cold},75\text{ per cent}}$  (95 per cent) and the  $\text{H}_2$   $h_{\text{H}_2,75\text{ per cent}}$  (95 per cent) are defined in a similar way. Results are averaged over time from  $t = 250$  Myr to 1 Gyr with a time interval of 1 Myr. The solid lines represent the average values in each radial bin and the filled areas show the  $\pm 1\sigma$  bands. The disc thickens at large radii. On the contrary, the cold atomic and molecular gas are confined within a thin layer in the mid-plane without flaring.



**Figure 13.** The radial profile of the surface density of the total gas ( $\Sigma_{\text{gas}}$ , black), star-forming gas ( $\Sigma_{\text{sfg}}$ , blue),  $\text{H}_2$  ( $\Sigma_{\text{H}_2}$ , red), star formation rate ( $\Sigma_{\text{SFR}}$ , green), mid-plane number density ( $n_0$ , cyan) and mid-plane pressure ( $P_0$ , brown). Results are averaged over time from  $t = 250$  Myr to 1 Gyr with a time interval of 1 Myr. The solid lines represent the average values in each radial bin and the filled areas show the  $\pm 0.5\sigma$  bands. The radial gradients of  $\Sigma_{\text{sfg}}$ ,  $\Sigma_{\text{H}_2}$  and  $\Sigma_{\text{SFR}}$  are much steeper than that of  $\Sigma_{\text{gas}}$ ,  $n_0$  and  $P_0$ , as a result of the decreasing cold gas fraction as  $R$  increases.

unshielded ISRF very quickly. Therefore, the  $\text{H}_2$  is also confined to the mid-plane.

#### 4.5.2 Radial profile

In Fig. 13, we show the radial profile of the surface density of the total gas ( $\Sigma_{\text{gas}}$ ), star-forming gas ( $\Sigma_{\text{sfg}}$ ),  $\text{H}_2$  ( $\Sigma_{\text{H}_2}$ ), star formation rate ( $\Sigma_{\text{SFR}}$ ), as well as the mid-plane number density ( $n_0$ ) and mid-plane pressure ( $P_0$ ), averaged over time from  $t = 250$  Myr to 1 Gyr with a time interval of 1 Myr. The solid lines are the arithmetic average and the filled areas are  $\pm 0.5\sigma$  bands in the given radial bins ( $\pm 1\sigma$  bands become negative in some bins and thus cannot be shown in log-scale). The mid-plane is defined as a thin layer within  $|z| < 0.1$  kpc.

An interesting feature in Fig. 13 is that  $\Sigma_{\text{SFR}}$  drops much faster than  $\Sigma_{\text{gas}}$  as  $R$  increases, i.e. the gas depletion time  $t_{\text{dep}} \equiv \Sigma_{\text{gas}}/\Sigma_{\text{SFR}}$  increases rapidly as  $R$  increases. Elmegreen & Hunter (2015) explored a possible explanation that the flaring of the disc (cf. Fig. 12) makes the density of the mid-plane, where star formation takes place, drop faster than the projected surface density. Our re-

sults do not favour this explanation as the radial gradient of  $n_0$  and  $\Sigma_{\text{gas}}$  do not differ much. Therefore, disc flaring is not responsible for the gradient difference between  $\Sigma_{\text{SFR}}$  and  $\Sigma_{\text{gas}}$ . Since the majority of the gas is in the warm phase of  $T \approx 10^4$  K, the mid-plane pressure  $P_0$  also shows a similar radial gradient as that of  $n_0$ . On the other hand,  $\Sigma_{\text{SFR}}$ ,  $\Sigma_{\text{sfg}}$  and  $\Sigma_{\text{H}_2}$  all show much steeper radial gradients. Roughly speaking, these three quantities all trace the cold gas, while  $\Sigma_{\text{gas}}$ ,  $n_0$  and  $P_0$  trace the total gas. Therefore, we can infer that the gradient difference between the former and the latter stems from the fact that the cold gas fraction declines as  $R$  increases.

The radius where the cold gas fraction drops to zero (as do  $\Sigma_{\text{SFR}}$ ,  $\Sigma_{\text{sfg}}$  and  $\Sigma_{\text{H}_2}$ ), denoted as  $R_{\text{cool}}$ , is about 1.4, 2.1 and 1.7 kpc in *G1D01*, *G1D001* and *G01D01*, respectively. The mid-plane density  $n_0$  at  $R_{\text{cool}}$  roughly corresponds to  $n_{\text{cool}}$ , the maximum density that keeps the equilibrium temperature at about  $10^4$  K (cf. Fig. 2). The gas is unable to cool below  $10^4$  K at  $R > R_{\text{cool}}$  because the mid-plane density is too low to trigger thermal-gravitational instability, which leads to an abrupt cutoff of  $\Sigma_{\text{SFR}}$ ,  $\Sigma_{\text{sfg}}$  and  $\Sigma_{\text{H}_2}$ . SN feedback is expected to scale with  $\Sigma_{\text{SFR}}$  and is therefore insignificant around  $R_{\text{cool}}$  compared to photoelectric heating, which does not

vary with  $\Sigma_{\text{SFR}}$  in our model. As a result, although star formation in the inner region is regulated by the SN feedback,  $R_{\text{cool}}$  is controlled by the photoelectric heating and thus varies with  $G_0$  and  $D$ . This is a consequence of our oversimplified assumption of spatially constant  $G_0$ .

In the inner part of the disc, the values of  $\Sigma_{\text{SFR}}$  and  $\Sigma_{\text{sfg}}$  are similar for the three runs, while  $\Sigma_{\text{H}_2}$  differs dramatically, demonstrating again that  $H_2$  formation is much more sensitive to variations of  $G_0$  and  $D$  than star formation is. The excellent match between  $\Sigma_{\text{sfg}}$  and  $\Sigma_{\text{H}_2}$  in *G01D01* is coincidental. As shown in both Figs 10 and 11, the star-forming gas is not necessarily  $H_2$ -dominated and vice versa.

#### 4.6 Density distribution function

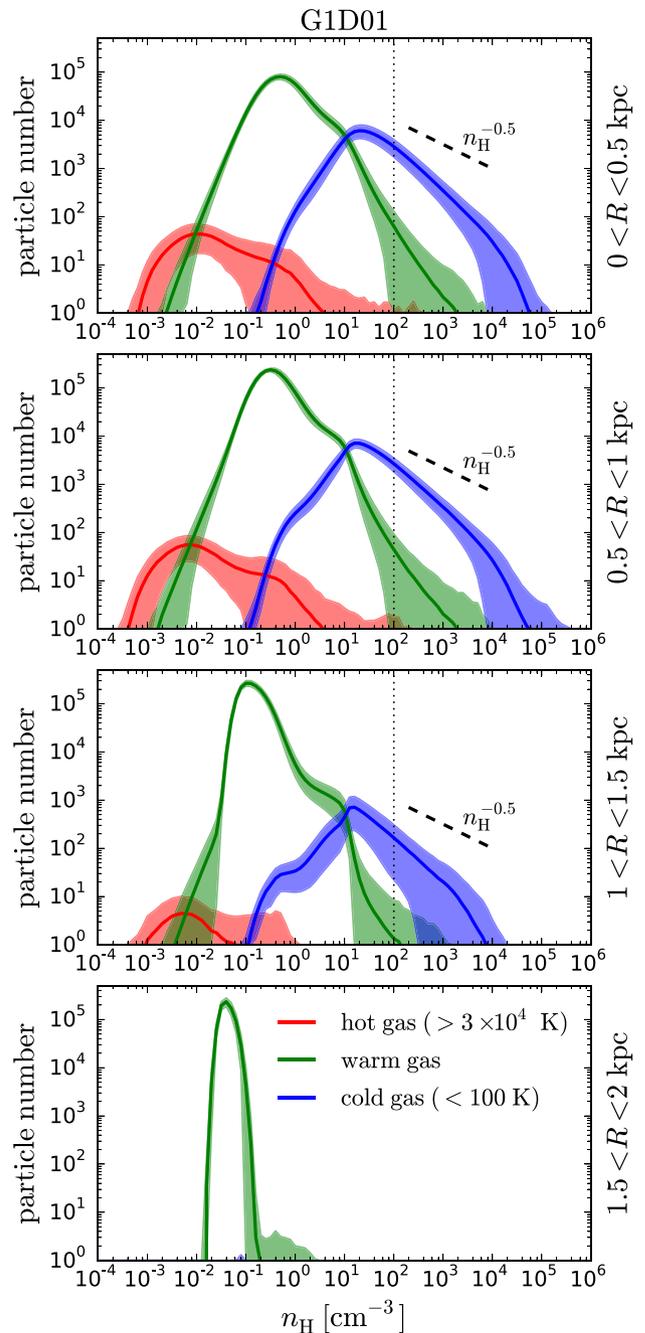
In Fig. 14, we show the density distribution function  $N(n_{\text{H}})$  of the hot ( $T > 3 \times 10^4$  K, in red), warm ( $100 < T < 3 \times 10^4$  K, in green) and cold ( $T < 100$  K, in blue) gas within the mid-plane  $|z| < 0.2$  kpc in four different radial regions in *G1D01*, averaged over time from  $t = 250$  Myr to 1 Gyr with a time interval of 1 Myr. It is binned in the range of  $n_{\text{H}} = 10^{-5}$ – $10^6$   $\text{cm}^{-3}$  with 110 density bins in log-space. The solid curves are the average values and the filled areas are the  $\pm 1\sigma$  bands which indicate the fluctuations of the density distributions.

The cold gas fraction clearly decreases as  $R$  increases, in agreement with what can be inferred from Fig. 13. The hot gas fraction also decreases as  $R$  increases since the SN feedback occurs less frequently at large radii. At  $R > 1.5$  kpc, both the cold and hot gas are absent and the warm gas has a rather uniform (narrow) density distribution. The mean density of the warm gas decreases as  $R$  increases, which reflects the radial profile of the mid-plane density. On the other hand, the mean and the width of the cold gas density distribution do not change much with  $R$  (except in the last panel where there is no cold gas). This means that gas starts to cool efficiently above a certain density which is not a sensitive function of  $R$ . Since the mean density is higher in the inner part of the disc, a larger fraction of gas would have high enough density to cool and form cold gas, which explains the decreasing cold gas fraction as  $R$  increases.

Another feature is that the cold gas distribution above the star formation threshold density  $n_{\text{H,th}}$  (dotted lines) drops faster than  $n_{\text{H}}^{-0.5}$  (shown in dashed line segments). Since the SFR of a gas particle is  $\epsilon_{\text{sf}} m_{\text{gas}} / t_{\text{ff}} \propto n_{\text{H}}^{0.5}$ , gas particles with densities just above  $n_{\text{H,th}}$  would have a dominant contribution to the total star formation rate (SFR =  $\int_{n_{\text{H,th}}}^{\infty} \epsilon_{\text{sf}} m_{\text{gas}} / t_{\text{ff}}(n_{\text{H}}) N(n_{\text{H}}) dn_{\text{H}}$ ). This leads to a linear relation between the surface density of gas and SFR, with a depletion time roughly equals  $t_{\text{ff}}(n_{\text{H,th}}) / \epsilon_{\text{sf}}$  (see Section 4.7), even though the assumed relation is superlinear locally ( $\propto n_{\text{H}}^{1.5}$ ).

#### 4.7 Star formation

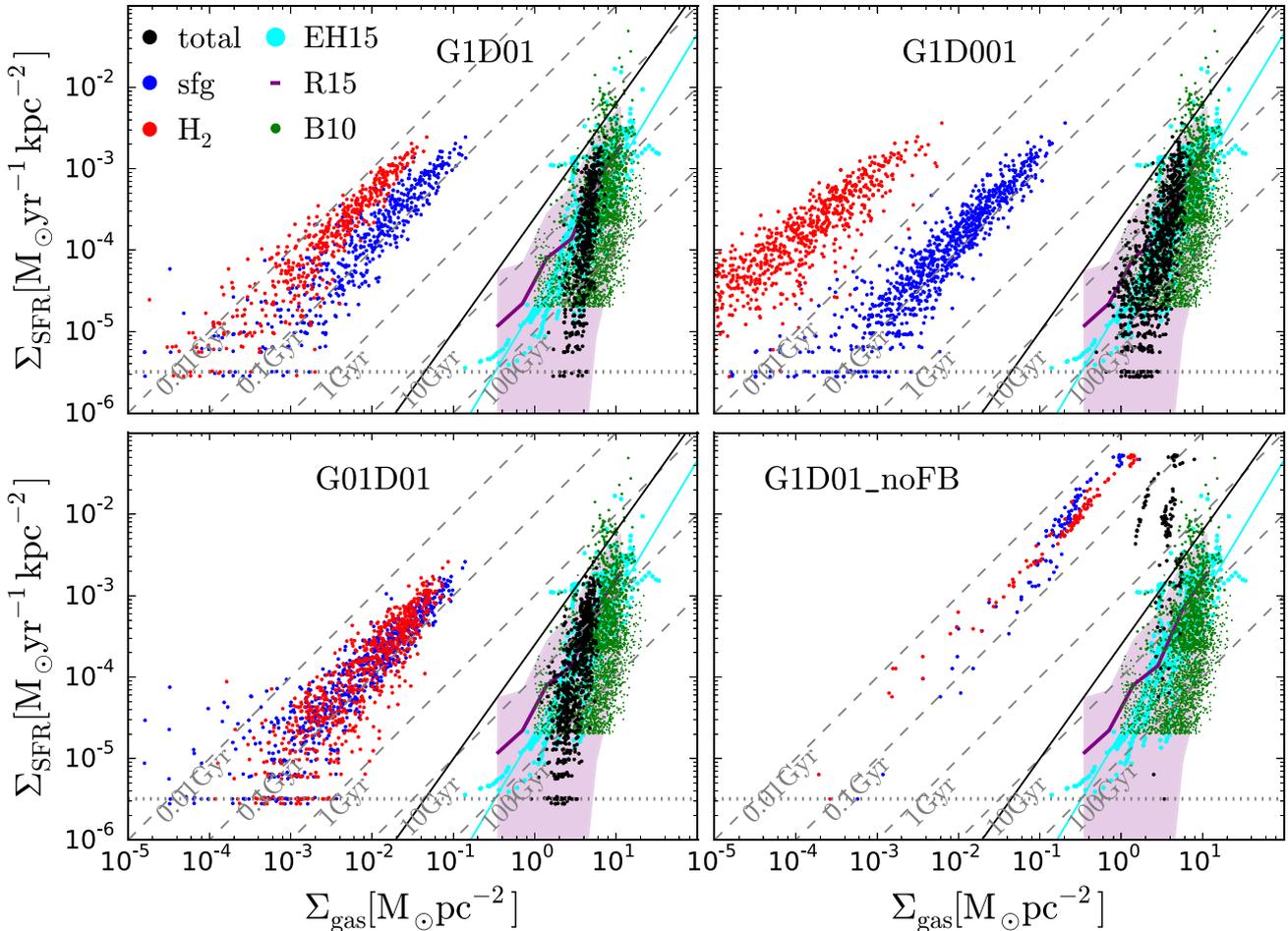
In Fig. 15, we show the SFR surface density  $\Sigma_{\text{SFR}}$  versus the surface density of the total gas ( $\Sigma_{\text{gas}}$ , in black), the star-forming gas ( $\Sigma_{\text{sfg}}$ , in blue) and the  $H_2$  ( $\Sigma_{\text{H}_2}$ , in red), respectively. All quantities are measured with aperture size  $l_{\text{ap}} = 500$  pc, as opposed to the radially averaged ones in Section 4.5.2. Each dot represents one aperture and the dots are the results from  $t = 250$  Myr to 1 Gyr with a time interval of 1 Myr (only a randomly chosen 4 per cent are shown for clarity). The solid line represents the conventional Kennicutt–Schmidt (KS) relation (Kennicutt 1998). The five dashed lines mark five different gas depletion times  $t_{\text{dep}} = 0.01, 0.1, 1, 10$  and 100 Gyr. The horizontal dashed lines show the lower limit for  $\Sigma_{\text{SFR}}$  for the



**Figure 14.** The density distribution of the hot ( $T > 3 \times 10^4$  K, in red), warm ( $100 < T < 3 \times 10^4$  K, in green) and cold ( $T < 100$  K, in blue) gas in *G1D01* in the regions of  $|z| < 0.2$  kpc and, from top to bottom,  $0 < R < 0.5$ ,  $0.5 < R < 1$ ,  $1 < R < 1.5$  and  $1.5 < R < 2$  kpc, respectively. The results are averaged over time from  $t = 250$  Myr to 1 Gyr with a time interval of 1 Myr. The solid lines represent the average values in each (log-)density bin and the filled areas show the  $\pm 1\sigma$  bands. The vertical dotted lines indicate the star formation threshold density and the dashed line segments show the slope of  $N \propto n_{\text{H}}^{-0.5}$  for reference.

given  $l_{\text{ap}}$  where one aperture only contains one star particle with age smaller than 5 Myr, i.e.  $\Sigma_{\text{SFR,min}} = m_{\text{star}} / (t_{\text{SF}} l_{\text{ap}}^2)$ .

For comparison with observations, the green dots are from Bigiel et al. (2010) with  $l_{\text{ap}} \approx 750$  pc. The purple line is the average value from the FIGGS survey from Roychowdhury et al. (2015) with  $l_{\text{ap}} \approx 400$  pc, while the filled area encloses the fifth and 95th percentile



**Figure 15.** SFR surface density versus surface density of the total (black), cold (blue) and  $\text{H}_2$  (red) gas, measured with aperture size  $l_{\text{ap}} = 500$  pc. The results from  $t = 250$  Myr to 1 Gyr with a time interval of 1 Myr are shown (only a randomly chosen 4 per cent are shown for clarity). The solid line represents the conventional KS relation and the dashed lines represent five different gas depletion times  $t_{\text{dep}} = 0.01, 0.1, 1, 10$  and 100 Gyr. The horizontal dotted lines are the minimum surface density allowed by our particle mass. The green and cyan dots are observational results from Bigiel et al. (2010) and Elmegreen & Hunter (2015), respectively. The cyan line is a fit to all the cyan points. The purple line is the average value from Roychowdhury et al. (2015) and the filled area encloses the fifth and 95th percentile. The steep slope in the total gas KS relation is due to the decreasing cold gas fraction at low gas surface density.

of the data. The cyan dots are the ring-averaged quantities and the cyan line is a fit to all the cyan points from Elmegreen & Hunter (2015).

The KS relation for the total gas surface density (black dots) lies below the black lines with a very steep slope, which is consistent with observations in this regime. On the other hand, the relation between  $\Sigma_{\text{sfg}}$  and  $\Sigma_{\text{SFR}}$  is almost linear, which implies a constant  $t_{\text{dep}}$  around 0.1 Gyr. As discussed in Section 4.6, gas with  $n_{\text{H}}$  just above the star formation threshold density  $n_{\text{H, th}} = 100 \text{ cm}^{-3}$  contributes the most to the total SFR. The depletion time of star-forming gas  $t_{\text{dep, sf}} \equiv \Sigma_{\text{sfg}} / \Sigma_{\text{SFR}}$  is therefore the free-fall time over the star formation efficiency  $t_{\text{ff}}(n_{\text{H, th}}) / \epsilon_{\text{sf}} \approx 0.1$  Gyr. Note that the value of  $t_{\text{dep, sf}}$  merely reflects our choice of  $n_{\text{H, th}}$  and therefore is not a prediction of our model. The fact that the  $\Sigma_{\text{sfg}} - \Sigma_{\text{SFR}}$  relation is approximately linear while the  $\Sigma_{\text{gas}} - \Sigma_{\text{SFR}}$  relation shows a much steeper slope is in fact already foreseeable in Fig. 13, as  $\Sigma_{\text{SFR}}$  and  $\Sigma_{\text{sfg}}$  show similar radial gradients much steeper than  $\Sigma_{\text{gas}}$ . The steep slope of the  $\Sigma_{\text{gas}} - \Sigma_{\text{SFR}}$  relation can be viewed as a dilution effect, as only a fraction of  $\Sigma_{\text{gas}}$  is actually participating in star formation.

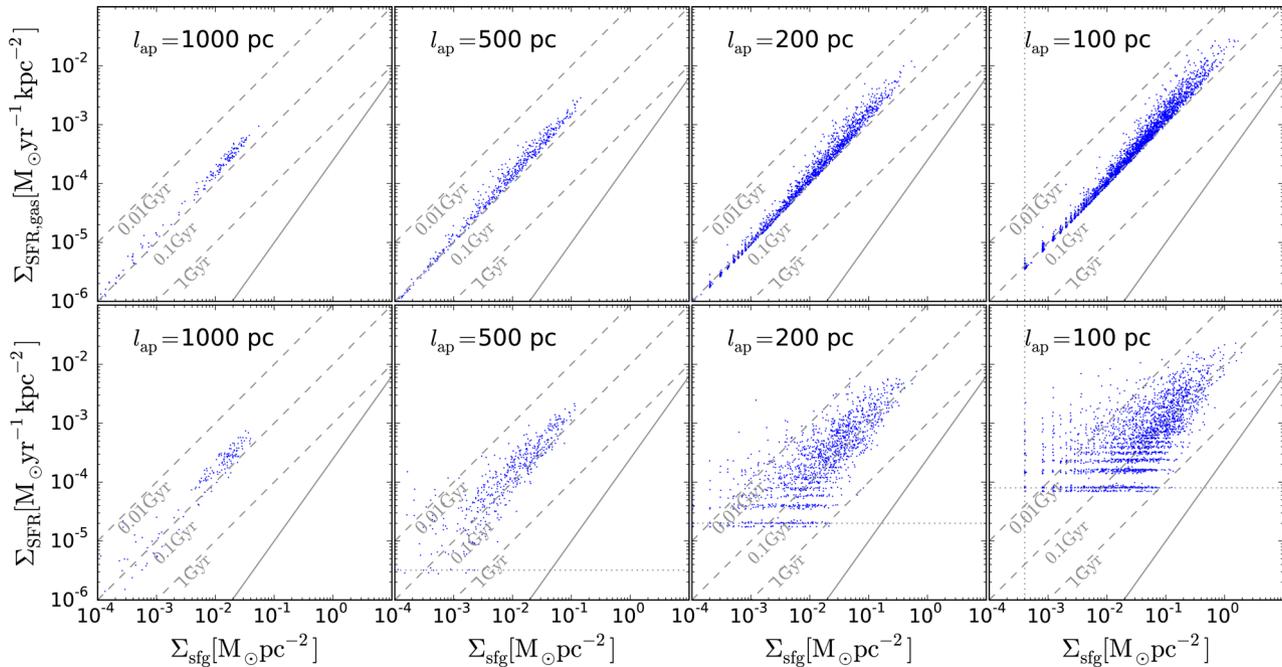
The relation between  $\Sigma_{\text{H}_2}$  and  $\Sigma_{\text{SFR}}$  is also approximately linear. However, defining the  $\text{H}_2$  depletion time,  $t_{\text{dep, H}_2} \equiv \Sigma_{\text{H}_2} / \Sigma_{\text{SFR}}$ , is not particularly useful and somewhat misleading, because the  $\text{H}_2$  is not

necessary for star formation, at least in our adopted star formation model. Moreover, the  $\text{H}_2$  is not a good tracer for the star-forming gas as shown in Figs 10 and 11. A small  $t_{\text{dep, H}_2}$  gives the wrong impression that the star formation in  $\text{H}_2$  is very efficient, though in reality it is simply due to the low  $\text{H}_2$  fraction (under the assumption that  $\text{H}_2$  is not a necessary condition for star formation).

Comparing *G1D01*, *G1D001* and *G01D01* in Fig. 15,  $\Sigma_{\text{H}_2}$  varies a lot horizontally while  $\Sigma_{\text{sfg}}$  is almost unchanged. Again, this indicates that the  $\text{H}_2$  formation is much more sensitive to the variation of  $G_0$  and  $D$  than the thermal properties of the gas (and hence star formation). In the no-feedback run *G1D01\_noFB*, the total gas KS relation is orders of magnitude above the solid line, which is obviously inconsistent with observations in this regime.

#### 4.7.1 The effect of aperture size

In Fig. 16, we show the relation between  $\Sigma_{\text{sfg}}$  and  $\Sigma_{\text{SFR}}$  in *G1D01* with four different aperture sizes ( $l_{\text{ap}} = 1000, 500, 200$  and 100 pc from left to right). Results from  $t = 250$  Myr to 1 Gyr with a time interval of 1 Myr are overplotted in each panel (only a randomly chosen 4 per cent are shown for clarity). The y-axis in the top row  $\Sigma_{\text{SFR, gas}}$  is the instantaneous SFR of the gas particles, while in the



**Figure 16.** The relation between  $\Sigma_{\text{sfg}}$  and  $\Sigma_{\text{SFR}}$  in *GID01* with four different aperture sizes ( $l_{\text{ap}} = 1000, 500, 200$  and  $100$  pc from left to right). Only a randomly chosen 4 per cent are shown for clarity. In the top panels  $\Sigma_{\text{SFR,gas}}$  are the instantaneous SFR from the gas particles, while in the bottom panels  $\Sigma_{\text{SFR}}$  are obtained by counting the young stars formed in the last 5 Myr. The vertical and horizontal dotted lines are the minimum surface density allowed by particle mass.

bottom row  $\Sigma_{\text{SFR}}$  is the SFR calculated by counting newly formed (age  $< 5$  Myr) stars (see Section 2.5). The horizontal and vertical dashed lines show the lower limit for  $\Sigma_{\text{SFR}}$  and  $\Sigma_{\text{sfg}}$ , respectively, for the given  $l_{\text{ap}}$ .

The top row of Fig. 16 shows a very tight  $\Sigma_{\text{sfg}} - \Sigma_{\text{SFR,gas}}$  relation that is very close to linear. The relation remains tight when  $l_{\text{ap}}$  shrinks. This suggests that the underlying cold gas density distribution does not change much from pixel to pixel even at the scale of 100 pc.

As discussed in Section 2.5, the instantaneous SFR of gas is not an observable and is only a conveniently defined input quantity for our model to estimate the future star formation activities. If we look at how many young stars really formed in the simulations (the bottom row of Fig. 16), which is observable, the relation becomes less tight, with a trend of increased scatter at lower  $\Sigma_{\text{sfg}}$ . The overall scatter increases significantly as  $l_{\text{ap}}$  shrinks. This is due to the time evolution effect of the star formation process that is discussed in detail in Schrubba et al. (2010) and Kruijssen & Longmore (2014). This phenomenon would not be visible if we were to look at the instantaneous SFR.

## 5 DISCUSSION

### 5.1 Non-equilibrium $H_2$ formation

In dwarf galaxies, the time-scale for  $H_2$  formation is longer than that in normal spiral galaxies because both the densities and the DGR are lower (equation 10). Due to the long chemical time-scales compared to the dynamical times, the gas is far out of chemical equilibrium locally (cf. Fig. 10). It is therefore essential to incorporate a time-dependent chemistry network in simulations to correctly predict the spatial distribution of  $H_2$ . On galactic scales, as there are both significant amounts of  $H_2$  over- and underabundant gas, the total  $H_2$

mass is only slightly overpredicted by the equilibrium prediction (cf. Fig. 6). However,  $F_{H_2,\text{eq}}$  shown in Fig. 6 is calculated using the self-shielding factor that is obtained from the non-equilibrium  $f_{H_2}$ . If one uses an equilibrium recipe in simulations, the  $H_2$  fraction is likely to be even more severely overestimated than in Fig. 6 because of the non-linearity that enters equation (14) via  $f_{\text{sh}}$ . Therefore, a time-dependent chemistry network is still necessary even if one is merely interested in the total  $H_2$  mass in a galaxy.

### 5.2 Star formation in $H\text{I}$ -dominated gas

Figs 10 and 11 demonstrate that the reservoir for star formation is not necessarily  $H_2$ -dominated (and vice versa), which means that the correlation between star formation and  $H_2$  breaks down in our simulated dwarf galaxy. This is in line with Glover & Clark (2012c) who simulated an isolated  $10^4 M_{\odot}$  cloud with much higher resolution ( $m_{\text{gas}} = 5 \times 10^{-3} M_{\odot}$ ,  $N_{\text{ngb}} = 100$ ). On the other hand, Krumholz (2012) compared the time-scales of  $H_2$  formation, radiative cooling and gravitational free-fall, and predicted the correlation would break down only when the metallicity is below a few per cent of  $Z_{\odot}$ , which is much lower than  $Z = 0.1 Z_{\odot}$  adopted in this work. One obvious reason for the discrepancy is the assumption of a linear  $Z$ -DGR relation in Krumholz (2012), as  $R_{\text{form}}$  mainly depends on  $D$  and not on  $Z$  per se. However, the star-forming gas is  $H\text{I}$ -dominated even in *GID01* and *G01D01* where we do assume a linear  $Z$ -DGR relation. One possible explanation is that feedback-driven turbulence constantly disrupts high-density clouds where most  $H_2$  formation takes place. Instead, the star-forming gas becomes  $H_2$ -dominated if we switch off stellar feedback (*GID01\_noFB*). In addition, Krumholz (2012) boosted the  $H_2$  formation rate coefficient,  $R_{\text{form}}$ , by a clumping factor  $C = 10$  to account for the density distribution below the scale the density is defined. The value of  $C$  is however very uncertain, and the prescription that simply takes it as a multiplication

factor for  $R_{\text{form}}$  is likely to overpredict the  $\text{H}_2$  formation rate (Micic et al. 2012). On the other hand, our simulations do not adopt the clumping factor (or  $C = 1$ , equivalently) and thus may underestimate the  $\text{H}_2$  fraction if there is significant clumping below our resolution scales. In the turbulent ISM picture, the gas is thought to be clumpier at large scales and more uniform at small scales. Mac Low & Glover (2012) conducted high-resolution ( $\approx 0.1$  pc) turbulence simulations and reported an effective clumping factor  $C \approx 2$  for turbulent clouds at the length scale of 10 pc. Since our effective spatial resolution is about 2 pc, we do not expect that the  $\text{H}_2$  fraction in our simulations would be severely underestimated.

In massive spiral galaxies, the use of  $\text{H}_2$ -dependent star formation recipes in hydrodynamical simulations is justified (though probably unnecessary) since most star-forming gas is also expected to be  $\text{H}_2$ -dominated. In dwarf galaxies, however, the star-forming gas is not necessarily  $\text{H}_2$ -dominated and vice versa. The  $\text{H}_2$ -dependent star formation recipes would therefore be inappropriate. This is especially true when the DGR falls below the linear  $Z$ -DGR relation. The SFR in our simulated galaxy would be much lower if we restricted star formation to occur only in  $\text{H}_2$ .<sup>4</sup>

### 5.3 The KS relation for dwarf galaxies

Observations of nearby star-forming galaxies have shown a roughly constant gas depletion time  $t_{\text{dep}} \approx 1$  Gyr at  $\Sigma_{\text{gas}} \gtrsim 10 M_{\odot} \text{pc}^{-2}$ , and a clear transition to much longer  $t_{\text{dep}}$  at  $\Sigma_{\text{gas}} \lesssim 10 M_{\odot} \text{pc}^{-2}$ , the regimes where most dwarf galaxies reside (see e.g. Bigiel et al. 2008, 2010; Schruba et al. 2011). The transition leads to a steepening of the slope in the KS relation, often referred to as the star formation threshold (not to be confused with  $n_{\text{H, th}}$ ). Our simulations suggest that such a transition does not correspond to the  $\text{H}_2$ -to- $\text{H I}$  transition, as the  $\text{H}_2$  fraction differs a lot while the KS relation for the total gas remains almost unchanged in different runs with feedback. Instead, this is a consequence of the dilution effect, as only a fraction of gas is cold and dense enough to form stars. Even though the mid-plane density decreases at almost the same rate as  $\Sigma_{\text{gas}}$ , the cold gas fraction also decreases with  $\Sigma_{\text{gas}}$  because a smaller fraction of gas would be dense enough to cool. If the density distribution of the cold gas does not change its shape but only its normalization factor from pixel to pixel, the  $\Sigma_{\text{sf}} - \Sigma_{\text{SFR}}$  relation should be linear, and so the  $\Sigma_{\text{gas}} - \Sigma_{\text{SFR}}$  would be superlinear. This dilution effect of the warm gas would not be visible at higher  $\Sigma_{\text{gas}}$  where the system is already cold-gas dominated (by mass), as is the case in most spiral galaxies. Our simulations do not cover high enough  $\Sigma_{\text{gas}}$  to see the transition, but we can speculate that the transition from a linear KS relation to a much steeper slope corresponds to a transition from a cold-gas dominated system to a warm-gas dominated one (by mass).

On the other hand, the slopes of the  $\Sigma_{\text{gas}} - \Sigma_{\text{SFR}}$  relation shown in Fig. 15 are probably steeper than observed due to our simplistic assumption of constant  $G_0$ . Physically, the strength of the diffuse ISRF is expected to scale with the average SFR, and therefore should also decrease as  $R$  increases until it drops to the level of the extragalactic UV background. However, it is unclear what spatial

scales are appropriate to define the averaged SFR that contributes to the local  $G_0$ . Our assumption of constant  $G_0$  is likely to be a severe overestimation at large radii. As discussed in Section 4.5.2, this leads to an abrupt truncation at  $R_{\text{cool}}$ , which in turn leads to very steep slopes in Fig. 15. In reality,  $G_0$  should decrease as  $R$  increases, and the cold gas fraction would not drop as abruptly as  $R$  increases. This is because while the mean density is lower at large  $R$ ,  $n_{\text{cool}}$  is also lower as photoelectric heating is weaker. The gas would therefore still be able to cool below  $10^4$  K at much larger radii up to the point where the gas is in thermal equilibrium at  $10^4$  K due to the extragalactic UV background (Schaye 2004) or it has a cooling time longer than the Hubble time. As such, the slope in the KS relation should be flatter than in Fig. 15. Note that the slope in Fig. 15 is slightly flatter for the runs with less efficient photoelectric heating, as the SN feedback which controls the cold gas fraction in these runs naturally scales with  $R$  in the inner region. Our result is consistent with Richings & Schaye (2016) who simulated slightly more massive isolated galaxies with a more sophisticated chemistry network and found that a weaker ISRF leads to a flatter slope and a lower threshold surface density. It is also in line with Bigiel et al. (2010) where they suggested an ‘S-shape’ KS relation where the slope flattens at disc outskirts. Interestingly, Roychowdhury et al. (2015) reported a much flatter slope of 1.5 for the average SFR in a given gas surface density bin, arguing that the individual measurement can be severely biased at low surface densities due to the stochasticity of star formation and thus only the average SFR can be faithfully obtained.

### 5.4 Spatial variations of ISRF and DGR

Our assumption of constant  $G_0$  and  $D$  is obviously an oversimplification. As already discussed,  $G_0$  is expected to scale with the averaged SFR. Moreover, at smaller scales in the vicinity of the young stars, the radiation field would easily be orders of magnitude stronger than the diffuse background, forming the photon-dominated regions around young stars which is not included in our model. Although the radiation field would drop rapidly away from the sources (geometrical effect), it can still potentially destroy a substantial amount of  $\text{H}_2$ , which makes the  $\text{H}_2$  fraction even lower. A more sophisticated radiative transfer is required to properly model the small-scale spatial variations of  $G_0$ .

On the other hand, the assumption of a constant DGR might actually be reasonable as the observed dust in our Galaxy seems to be well mixed with gas (Boulanger et al. 1996). Though locally dust is continuously created and destroyed via various physical processes (Zhukovska 2014; Feldmann 2015), this inhomogeneity can be effectively smoothed out through turbulent mixing. A dust evolution model is required to investigate such effects.

## 6 SUMMARY AND CONCLUSIONS

We have conducted high-resolution ( $m_{\text{gas}} = 4 M_{\odot}$ ,  $N_{\text{ngb}} = 100$ ) hydrodynamical simulations for an isolated star-forming dwarf galaxy run for 1 Gyr. Our model includes self-gravity, non-equilibrium cooling and chemistry, shielding from a uniform and constant ISRF, star formation, stellar feedback and metal enrichment self-consistently. We have investigated the physical properties of the ISM in low-metallicity environments on galactic scales. Our main results can be summarized as follows.

<sup>4</sup> More precisely, if we suddenly switch to a  $\text{H}_2$ -dependent star formation model during our simulations, the SFR would be much lower for a short period of time, which could lead to further collapse of the gas (but see Pelupessy & Papadopoulos 2009). The system may still eventually reach a new quasi-equilibrium state with a much clumpier density structure that cannot be resolved with our resolution.

(i) The ISM in dwarf galaxies is dominated by the warm gas ( $100 \text{ K} < T \leq 3 \times 10^4 \text{ K}$ ) both in mass and in volume (Fig. 7). The hot gas ( $T > 3 \times 10^4 \text{ K}$ ) occupies about 10 per cent of volume but contributes little of the mass ( $\approx 0.1$  per cent). The cold gas ( $T < 100 \text{ K}$ ) contributes up to 1–10 per cent of the mass but occupies little volume ( $< 0.1$  per cent).

(ii) We found SN-driven galactic outflows in our simulated galaxy. The mass loading measured at  $|z| = \pm 2 \text{ kpc}$  is slightly larger than unity. A small fraction of outflowing gas at  $|z| = \pm 2 \text{ kpc}$  eventually escapes the halo. The outflowing gas is slightly enriched: the metallicity in the halo ( $Z_{\text{halo}}$ ) is about 20 per cent higher than that in the disc ( $Z_{\text{disc}}$ ).

(iii) The ISM in dwarf galaxies is a hostile environment for  $H_2$  formation because of its low DGR, strong ISRF and low densities in general. The  $H_2$  mass fraction of the galaxy is very low (less than 0.1 per cent) in all runs except for *GID01\_noFB* (Fig. 6). The  $H_2$  mass fraction is very sensitive to the variations of  $G_0$  and  $D$  while the gas thermal properties and hence the total SFR are insensitive to both  $G_0$  and  $D$ .

(iv) Feedback-driven turbulence keeps the gas from cooling below  $10^4 \text{ K}$  at  $n_H \lesssim 1 \text{ cm}^{-3}$  by reducing the local dynamical time to be smaller than the cooling time. Once the gas cools below  $10^4 \text{ K}$ , it would be shocked heated to  $10^4 \text{ K}$  in a dynamical time, and so the gas is out of thermal equilibrium in the presence of SN feedback (Fig. 9). SN feedback (which does not depend explicitly on both  $G_0$  and  $D$ ) therefore determines the gas distribution in the phase diagram and regulates star formation. However, the size of the star-forming region  $R_{\text{cool}}$  is controlled by the photoelectric heating (Fig. 13), as it controls the ability to cool and form stars in the first place.

(v) Since the chemical time-scales are much longer than the local dynamical times, the gas is far out of chemical equilibrium locally (Fig. 10). The star-forming gas is  $H I$  dominated in all runs except for *GID01\_noFB*. There is also a significant amount of  $H_2$ -dominated gas that resides in the diffuse and warm phase and that is not star forming (Fig. 11). This suggests that the correlation between  $H_2$  and star formation breaks down at low metallicity.

(vi) The scaleheight of the total gas increases with  $R$ . On the other hand, most of the cold gas ( $T < 100 \text{ K}$ ) and  $H_2$  are confined within a thin layer ( $|z| < 0.1 \text{ kpc}$ ) in the mid-plane with no sign of flaring (Fig. 12). However, the radial gradient of the mid-plane density  $n_0$  is only slightly steeper than  $\Sigma_{\text{gas}}$ , both of which are much shallower than the radial gradient of  $\Sigma_{\text{SFR}}$  (Fig. 13). Therefore, disc flaring does not account for the steep slopes in the  $\Sigma_{\text{gas}} - \Sigma_{\text{SFR}}$  relation (Fig. 15).

(vii) The cold gas fraction decreases with the mid-plane density because a lesser fraction of gas is dense enough to cool and form cold gas if the mid-plane (mean) density is lower. Meanwhile, the cold gas density distribution does not vary much from region to region, and so a constant fraction of the cold gas will be star forming. Therefore, the steep slopes in the KS relation are due to the dilution effect of warm gas: only a small fraction of gas is forming stars and this fraction decreases as  $\Sigma_{\text{gas}}$  decreases.

## ACKNOWLEDGEMENTS

We thank the referee for the useful and constructive comments on the paper. We thank Bruce Elmegreen and Sambit Roychowdhury for providing the observational data from Elmegreen & Hunter (2015) and Roychowdhury et al. (2015). TN acknowledges support from DFG priority program SPP 1573 ‘Physics of the interstellar medium’. SW acknowledges support by the DFG via SFB 956

‘Conditions and impact of star formation’ and SPP 1573, as well as by the Bonn-Cologne Graduate School. SCOG acknowledges financial support from the Deutsche Forschungsgemeinschaft via SFB 881, ‘The Milky Way System’ (sub-projects B1, B2 and B8) and SPP 1573, ‘Physics of the Interstellar Medium’ (grant number GL 668/2-1). SCOG further acknowledges support from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013) via the ERC Advanced Grant STARLIGHT (project number 339177).

## REFERENCES

- Agertz O. et al., 2007, MNRAS, 380, 963  
Aumer M., White S. D. M., Naab T., Scannapieco C., 2013, MNRAS, 434, 3142  
Bakes E. L. O., Tielens A. G. G. M., 1994, ApJ, 427, 822  
Barnes J., Hut P., 1986, Nature, 324, 446  
Bate M. R., Burkert A., 1997, MNRAS, 288, 1060  
Bergin E. A., Hartmann L. W., Raymond J. C., Ballesteros-Paredes J., 2004, ApJ, 612, 921  
Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846  
Bigiel F., Leroy A., Walter F., Blitz L., Brinks E., de Blok W. J. G., Madore B., 2010, AJ, 140, 1194  
Bigiel F. et al., 2011, ApJ, 730, L13  
Blondin J. M., Wright E. B., Borkowski K. J., Reynolds S. P., 1998, ApJ, 500, 342  
Bolatto A. D. et al., 2011, ApJ, 741, 12  
Bolatto A. D., Wolfire M., Leroy A. K., 2013, ARA&A, 51, 207  
Boulanger F., Abergel A., Bernard J.-P., Burton W. B., Desert F.-X., Hartmann D., Lagache G., Puget J.-L., 1996, A&A, 312, 256  
Christensen C., Quinn T., Governato F., Stilp A., Shen S., Wadsley J., 2012, MNRAS, 425, 3058  
Clark P. C., Glover S. C. O., Klessen R. S., 2012, MNRAS, 420, 745  
Cormier D. et al., 2015, A&A, 578, A53  
Cullen L., Dehnen W., 2010, MNRAS, 408, 669  
Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140  
Dehnen W., Aly H., 2012, MNRAS, 425, 1068  
Dobbs C. L., Glover S. C. O., Clark P. C., Klessen R. S., 2008, MNRAS, 389, 1097  
Draine B. T., Bertoldi F., 1996, ApJ, 468, 269  
Durier F., Dalla Vecchia C., 2012, MNRAS, 419, 465  
Elmegreen B. G., Hunter D. A., 2015, ApJ, 805, 145  
Feldmann R., 2015, MNRAS, 449, 3274  
Fu J., Guo Q., Kauffmann G., Krumholz M. R., 2010, MNRAS, 409, 515  
Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375  
Girichidis P. et al., 2016, MNRAS, 456, 3432  
Glover S. C. O., Clark P. C., 2012a, MNRAS, 421, 9  
Glover S. C. O., Clark P. C., 2012b, MNRAS, 421, 116  
Glover S. C. O., Clark P. C., 2012c, MNRAS, 426, 377  
Glover S. C. O., Clark P. C., 2014, MNRAS, 437, 9  
Glover S. C. O., Mac Low M.-M., 2007, ApJS, 169, 239  
Glover S. C. O., Mac Low M.-M., 2011, MNRAS, 412, 337  
Gnedin N. Y., Tassis K., Kravtsov A. V., 2009, ApJ, 697, 55  
Górski K. M., Hivon E., 2011, Astrophysics Source Code Library, record ascl:1107.018  
Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays, preprint (arXiv:astro-ph/0106018)  
Habing H. J., 1968, Bull. Astron. Inst. Neth., 19, 421  
Hollenbach D., McKee C. F., 1979, ApJS, 41, 555  
Hopkins P. F., 2013, MNRAS, 428, 2840  
Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950  
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581  
Hu C.-Y., Naab T., Walch S., Moster B. P., Oser L., 2014, MNRAS, 443, 1173

Hunter D. A. et al., 2012, *AJ*, 144, 134  
 Iwamoto K., Brachwitz F., Nomoto K., Kishimoto N., Umeda H., Hix W. R., Thielemann F.-K., 1999, *ApJS*, 125, 439  
 Karakas A. I., 2010, *MNRAS*, 403, 1413  
 Kennicutt R. C., Jr, 1998, *ApJ*, 498, 541  
 Kim C.-G., Ostriker E. C., 2015, *ApJ*, 802, 99  
 Kroupa P., 2001, *MNRAS*, 322, 231  
 Kruijssen J. M. D., Longmore S. N., 2014, *MNRAS*, 439, 3239  
 Krumholz M. R., 2012, *ApJ*, 759, 9  
 Krumholz M. R., McKee C. F., 2005, *ApJ*, 630, 250  
 Krumholz M. R., Leroy A. K., McKee C. F., 2011, *ApJ*, 731, 25  
 Kuhlen M., Krumholz M. R., Madau P., Smith B. D., Wise J., 2012, *ApJ*, 749, 36  
 Leroy A. K. et al., 2011, *ApJ*, 737, 12  
 Lucy L. B., 1977, *AJ*, 82, 1013  
 McKee C. F., Krumholz M. R., 2010, *ApJ*, 709, 308  
 Mac Low M.-M., Glover S. C. O., 2012, *ApJ*, 746, 135  
 Maoz D., Mannucci F., 2012, *PASA*, 29, 447  
 Michałowski M. J. et al., 2015, *A&A*, 582, A78  
 Micic M., Glover S. C. O., Federrath C., Klessen R. S., 2012, *MNRAS*, 421, 2531  
 Morris J. P., Monaghan J. J., 1997, *J. Comput. Phys.*, 136, 41  
 Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, 710, 903  
 Moster B. P., Naab T., White S. D. M., 2013, *MNRAS*, 428, 3121  
 Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493  
 Nelson R. P., Langer W. D., 1997, *ApJ*, 482, 796  
 Ostriker J. P., McKee C. F., 1988, *Rev. Mod. Phys.*, 60, 1  
 Ostriker E. C., McKee C. F., Leroy A. K., 2010, *ApJ*, 721, 975  
 Pelupessy F. I., Papadopoulos P. P., 2009, *ApJ*, 707, 954  
 Price D. J., 2008, *J. Comput. Phys.*, 227, 10040  
 Read J. I., Hayfield T., 2012, *MNRAS*, 422, 3037  
 Read J. I., Hayfield T., Agertz O., 2010, *MNRAS*, 405, 1513  
 Rémy-Ruyer A. et al., 2014, *A&A*, 563, A31  
 Renaud F. et al., 2013, *MNRAS*, 436, 1836  
 Richings A. J., Schaye J., 2016, *MNRAS*, 458, 270  
 Robertson B. E., Kravtsov A. V., 2008, *ApJ*, 680, 1083  
 Roychowdhury S., Huang M.-L., Kauffmann G., Wang J., Chengalur J. N., 2015, *MNRAS*, 449, 3700  
 Saitoh T. R., Makino J., 2009, *ApJ*, 697, L99  
 Saitoh T. R., Makino J., 2013, *ApJ*, 768, 44  
 Schaye J., 2004, *ApJ*, 609, 667  
 Schaye J. et al., 2015, *MNRAS*, 446, 521  
 Schrubba A., Leroy A. K., Walter F., Sandstrom K., Rosolowsky E., 2010, *ApJ*, 722, 1699  
 Schrubba A. et al., 2011, *AJ*, 142, 37  
 Schrubba A. et al., 2012, *AJ*, 143, 138  
 Sedov L. I., 1959, *Similarity and Dimensional Methods in Mechanics*. Academic Press, New York  
 Shetty R., Glover S. C., Dullemond C. P., Klessen R. S., 2011a, *MNRAS*, 412, 1686  
 Shetty R., Glover S. C., Dullemond C. P., Ostriker E. C., Harris A. I., Klessen R. S., 2011b, *MNRAS*, 415, 3253  
 Smith R. J., Glover S. C. O., Clark P. C., Klessen R. S., Springel V., 2014, *MNRAS*, 441, 1628  
 Somerville R. S., Popping G., Trager S. C., 2015, *MNRAS*, 453, 4337  
 Springel V., 2005, *MNRAS*, 364, 1105  
 Sternberg A., 1988, *ApJ*, 332, 400  
 Sternberg A., Le Petit F., Roueff E., Le Bourlot J., 2014, *ApJ*, 790, 10  
 Thompson R., Nagamine K., Jaacks J., Choi J.-H., 2014, *ApJ*, 780, 145  
 Tremonti C. A. et al., 2004, *ApJ*, 613, 898  
 Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518  
 Walch S., Naab T., 2015, *MNRAS*, 451, 2757  
 Walch S., Wünsch R., Burkert A., Glover S., Whitworth A., 2011, *ApJ*, 733, 47  
 Walch S. et al., 2015, *MNRAS*, 454, 238

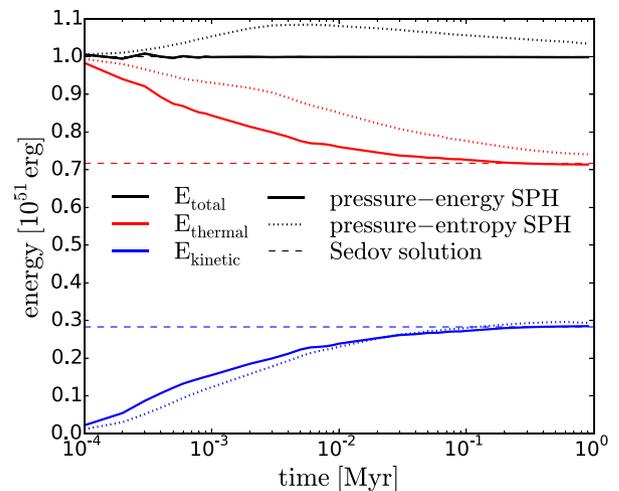
Wiersma R. P. C., Schaye J., Smith B. D., 2009, *MNRAS*, 393, 99  
 Wolfire M. G., Hollenbach D., McKee C. F., 2010, *ApJ*, 716, 1191  
 Woosley S. E., Weaver T. A., 1995, *ApJS*, 101, 181  
 Zhukovska S., 2014, *A&A*, 562, A76

## APPENDIX A: ENERGY CONSERVATION IN PRESSURE-BASED SPH

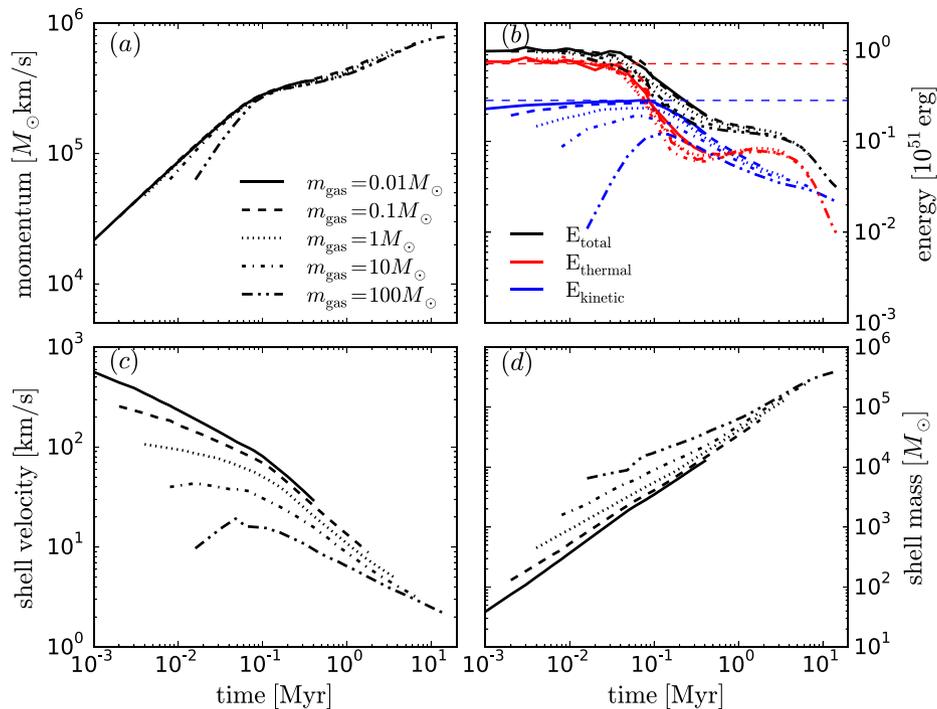
We investigate the energy conservation property of two different SPH formulations: pressure-entropy SPH and pressure-energy SPH (Hopkins 2013). We model a non-radiative blastwave in a uniform medium with the number density  $n = 100 \text{ cm}^{-3}$  and the initial temperature  $T = 1000 \text{ K}$ . A total energy of  $10^{51} \text{ erg}$  is injected into the neighbouring 100 particles in terms of thermal energy and is distributed by the smoothing kernel. The radiative cooling is turned off. A global timestep is used to ensure that the violation of energy conservation is not due to the adaptive timesteps. The particle mass is  $1 M_{\odot}$ .

In Fig. A1 we show the time evolution of the total energy (black), thermal energy (red) and kinetic energy (blue). The solid and dotted lines represent the results of using pressure-energy SPH and pressure-entropy SPH, respectively. The horizontal dashed lines are the exact solution for a non-radiative blastwave, the so-called Sedov solution (Sedov 1959): the total energy is conserved and the fraction of thermal and kinetic energy are about 73 and 27 per cent, respectively.

When using pressure-energy SPH, the total energy is conserved and the thermal and kinetic energy converge to the exact solution. On the other hand, when using pressure-entropy SPH, the total energy increases over time by up to about 10 per cent and then gradually decreases even if a global timestep is adopted. This is because when converting entropy into energy in pressure-entropy SPH, an estimate of density is required. There are two different ways to estimate the density: the entropy-weighted density  $\rho^e$  and the traditional mass-weighted density  $\rho^m$  (see e.g. Hu et al. 2014). In the Lagrangian formulation (Hopkins 2013),  $\rho^e$  is a natural choice for the density estimate. Choosing  $\rho^m$  leads to inconsistencies and so compromises the conservation property in dissipationless



**Figure A1.** Time evolution of the total energy (black), thermal energy (red) and kinetic energy (blue) for a non-radiative blastwave in a uniform medium. The solid and dotted lines represent the results of using pressure-energy SPH and pressure-entropy SPH, respectively. The horizontal dashed lines are the exact solution. Pressure-energy SPH conserves total energy and the thermal and kinetic energy converge to the exact solutions, while pressure-entropy SPH suffers from cumulative error of energy conservation.



**Figure B1.** Evolution of an SNR of energy  $E = 10^{51}$  erg with five different mass resolutions  $m_{\text{gas}} = 0.01$  (solid),  $0.1$  (dashed),  $1$  (dotted),  $10$  (dash-dotted) and  $100$  (double dot-dashed)  $M_{\odot}$  in a medium of number density  $n = 1 \text{ cm}^{-3}$ . Panel (a): the total radial momentum; panel (b): the total energy (black), thermal energy (red) and kinetic energy (blue), with the dashed lines showing the energy partition (73 per cent thermal and 27 per cent kinetic) in the Sedov phase; panel (c): the shell velocity, defined as the total momentum divided by the shell mass; panel (d): the shell mass. The SNR is defined by all particles whose velocity  $> 0.1 \text{ km s}^{-1}$ .

systems. However, when dissipation (artificial viscosity) is included,  $\rho^e$  causes a large error at entropy discontinuities due to its entropy-weighting and thus also violates energy conservation (Hu et al. 2014). This leads to a dilemma when converting entropy into energy in pressure-entropy SPH. Therefore, in this work we adopt pressure-energy SPH which shows much better energy conservation property.

## APPENDIX B: RESOLUTION STUDY OF THE SN FEEDBACK

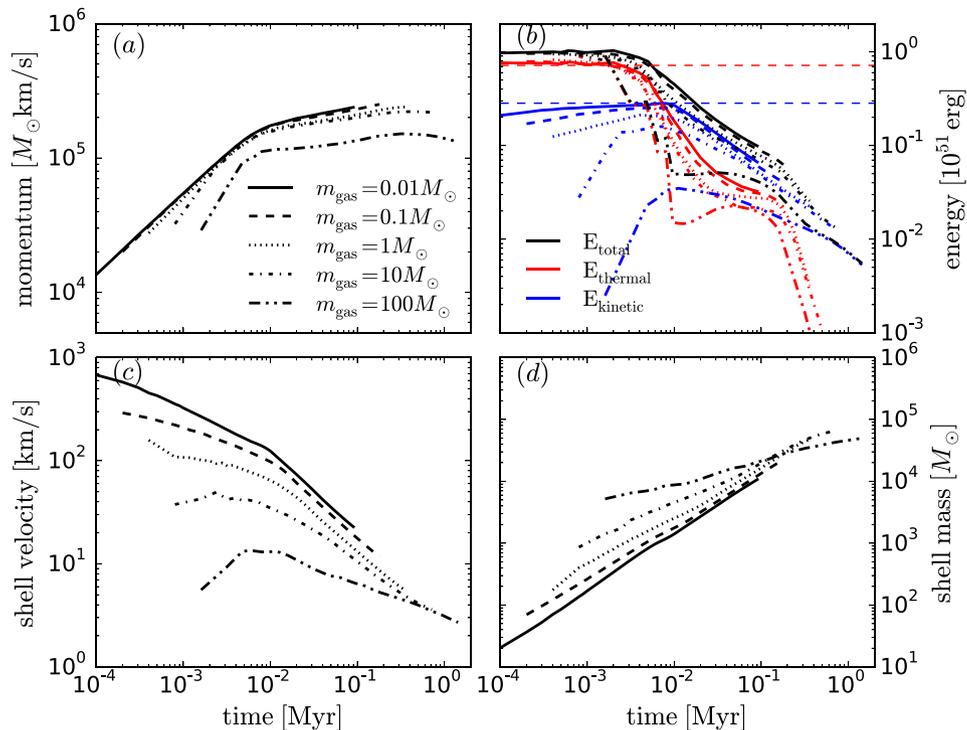
In order to investigate what resolution is required to fully resolve the SN feedback without suffering from the numerical overcooling problem, we perform a resolution study of a single SN explosion in a uniform medium. A total energy of  $10^{51}$  erg is injected into the neighbouring 100 particles in terms of thermal energy and is distributed by the smoothing kernel. The medium has similar ISM properties as our fiducial run in Section 4 with initial temperature  $T = 1000$  K. In Fig. B1, we show the time evolution of the SN feedback with five runs with different particle mass ( $0.01$ ,  $0.1$ ,  $1$ ,  $10$  and  $100 M_{\odot}$ ). Panel (a) shows the total linear momentum and panel (b) shows the total energy (in black), thermal energy (in red) and kinetic energy (in blue). Panels (c) and (d) show the shell velocity and shell mass, respectively. The shell mass is defined as the total mass of particles whose velocity  $v > 0.1 \text{ km s}^{-1}$ , i.e. the shock-accelerated particles. The shell velocity is defined as the total momentum divided by the shell mass. The medium number density  $n = 1 \text{ cm}^{-3}$ .

In the initial non-radiative phase (the so-called Sedov phase) the total energy is conserved, and the fraction of thermal and kinetic energy will be about 73 and 27 per cent, respectively (see e.g. Walch

& Naab 2015), as we show by the two horizontal dashed lines. If the kinetic energy remains constant, then as more material is swept up, the total mass increases and the system gains momentum while decreasing its velocity.

The total energy is no longer conserved once the shell material starts to cool ( $t \approx 0.07$  Myr) and loses thermal energy. In the limit of the complete removal of thermal energy, the shock will enter a momentum-conserving phase where the advance of the shock relies entirely on the inertia of the shell. However, most of the shell material would first cool rapidly down to  $T \lesssim 10^4$  K where the cooling rate drops significantly below the Lyman  $\alpha$  peak, and then cools much more gradually afterwards. Therefore, there is first a rapid drop in the total energy ( $t \approx 0.07$ – $0.5$  Myr) followed by a much more gradual one ( $t \approx 0.5$ – $5$  Myr). There is also a slight increase in thermal energy between  $t \approx 0.5$  and  $5$  Myr as the conversion from kinetic to thermal energy counteracts the cooling. This residual thermal energy provides extra fuel for further momentum gain, though not as efficient as in the Sedov phase.

The numerical overcooling problem can be clearly seen in the energy evolution. At worse resolution, the conversion from thermal to kinetic energy becomes slower. If the system starts to cool before the blastwave is fully developed, the kinetic energy will be underestimated. It is interesting to note that the total momentum is actually not as sensitive to resolution. Even at our worst resolution the momentum evolution seems to agree with high-resolution runs well. However, the swept-up mass differs quite dramatically. The low-resolution runs seem to sweep up too much mass due to their inability of resolving the thin shell. The overestimated mass gives rise to an underestimated velocity and kinetic energy, even though the total momentum generation is approximately correct.



**Figure B2.** Same as Fig. B1 but in a medium of number density  $n = 100 \text{ cm}^{-3}$ .

Fig. B2 shows another test similar to Fig. B1 except in a denser medium of  $n = 100 \text{ cm}^{-3}$ . The cooling time  $t_{\text{cool}}$  is about 10 times shorter than in the  $n = 1 \text{ cm}^{-3}$  medium, which is consistent with the scaling  $t_{\text{cool}} \propto n^{-9/17}$  reported in Blondin et al. (1998). The general features are quite similar to the low-density case. The momentum again shows only weak dependence on resolution, except for the lowest resolution run, where the momentum is underestimated by about a factor of 2. The energy (both kinetic and thermal) evolution shows a slightly more underestimation for given resolution, suggesting that better resolution is required to avoid numerical overcooling in denser medium.

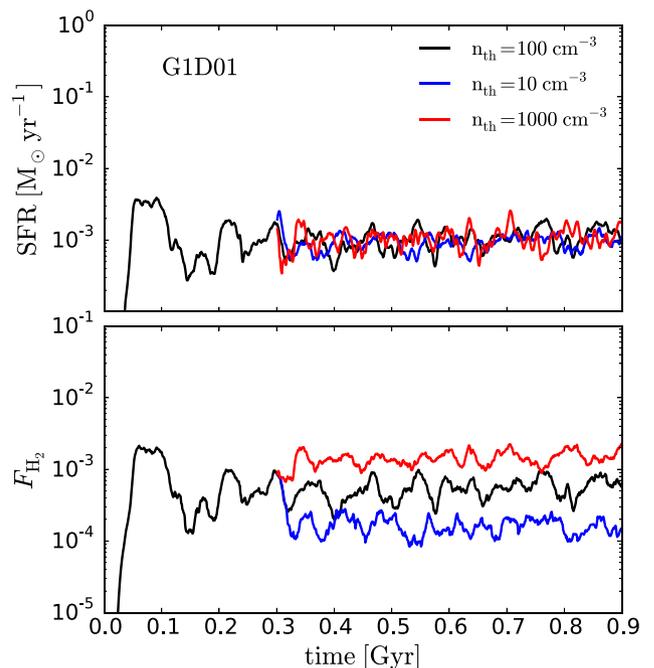
In summary, the results suggest that for a  $10^{51}$  erg SN explosion, it requires a resolution of about  $1 M_{\odot}$  to recover the reasonably converged velocities. In the unresolved case where too much mass is swept up, the ISM would only be accelerated to a much underestimated velocity, even if the right amount of momentum were generated. This has a direct impact on the capability of driving galactic outflows as it is the outflowing velocity that determines whether a cloud is able to escape the potential well.

## APPENDIX C: PARAMETER DEPENDENCE

We explore the effects of varying the values of the physical parameters in our models. Instead of starting from the initial conditions described in Section 4.1, all runs in this section are started from the snapshot at  $t = 300 \text{ Myr}$  of the *G1D01* run.

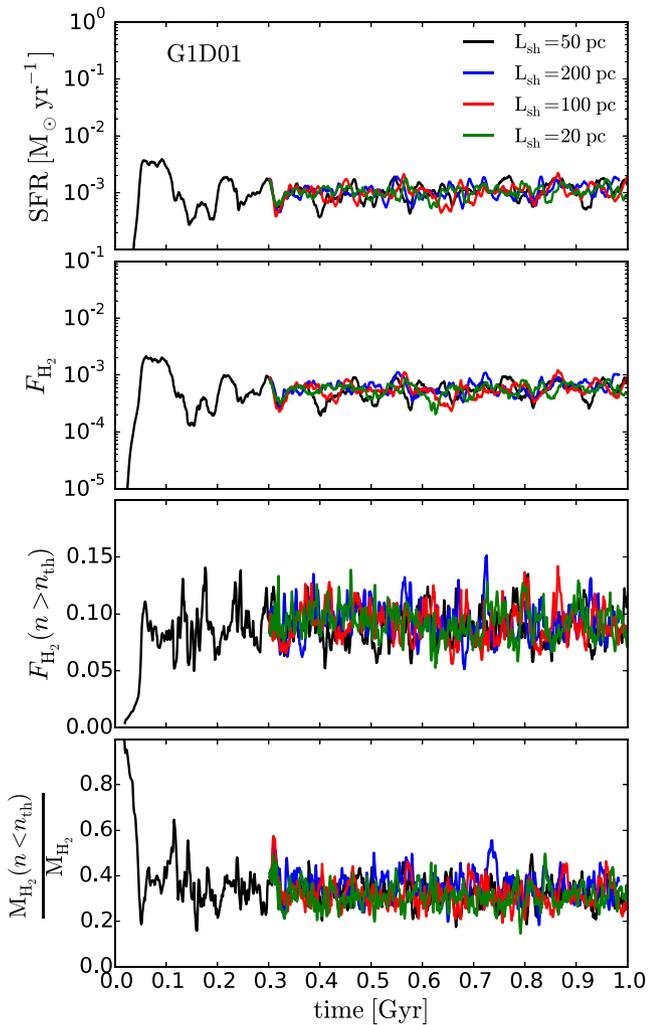
### C1 Star formation threshold density

Fig. C1 shows the time evolution of the total SFR (*upper panel*) and the  $\text{H}_2$  mass fraction (*lower panel*) in the ISM ( $R < 2 \text{ kpc}$



**Figure C1.** Time evolution of the total SFR (*upper panel*) and the  $\text{H}_2$  mass fraction (*lower panel*) in the ISM ( $R < 2 \text{ kpc}$  and  $|z| < 1 \text{ kpc}$ ) with the *G1D01* model, with the star formation threshold density  $n_{\text{th}} = 100$  (black), 10 (blue) and  $1000 \text{ cm}^{-3}$  (red).

and  $|z| < 1 \text{ kpc}$ ), with the star formation threshold density  $n_{\text{th}} = 100$  (black), 10 (blue) and  $1000 \text{ cm}^{-3}$  (red), respectively. While the SFR is insensitive to the choice of  $n_{\text{th}}$ , the  $\text{H}_2$  mass fraction does show significant differences with different  $n_{\text{th}}$ , as a higher  $n_{\text{th}}$  leads



**Figure C2.** Time evolution of the total SFR (first panel), the  $H_2$  mass fraction (second panel), the  $H_2$  mass fraction for the dense ( $n > n_{\text{th}}$ ) gas (third panel), and the fraction of the  $H_2$  mass that resides in the diffuse ( $n < n_{\text{th}}$ ) gas (fourth panel) in the ISM ( $R < 2$  kpc and  $|z| < 1$  kpc) with the G1D01 model. None of these quantities are sensitive to the choice of  $L_{\text{sh}}$ .

to a more clumpy ISM structure (broader density distribution) and more dense gas which increases the  $H_2$  mass fraction. However, even in the  $n_{\text{th}} = 1000 \text{ cm}^{-3}$  run, the  $H_2$  mass fraction remains very small,  $F_{H_2} \sim 10^{-3}$ . We note that with  $n_{\text{th}} = 1000 \text{ cm}^{-3}$  the star-forming gas would be strongly unresolved and therefore is not an appropriate choice for our resolution. Setting  $n_{\text{th}} = 10 \text{ cm}^{-3}$  makes the system even more  $H_2$ -poor, though the star formation efficiency  $\epsilon = 2$  per cent may be too high for such a choice.

## C2 Shielding length

Fig. C2 shows the time evolution of the total SFR (*first panel*), the  $H_2$  mass fraction (*second panel*), the  $H_2$  mass fraction for the dense ( $n > n_{\text{th}}$ ) gas (*third panel*), and the fraction of the  $H_2$  mass that resides in the diffuse ( $n < n_{\text{th}}$ ) gas (*fourth panel*) in the ISM ( $R < 2$  kpc and  $|z| < 1$  kpc). Different lines represent different choices of the shielding length  $L_{\text{sh}} = 20$  pc (green), 50 pc (black), 100 pc (red), and 200 pc (blue), respectively. None of these quantities are sensitive to the choice of  $L_{\text{sh}}$  (see the discussion in Section 2.4).

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