

# Positive feedback at the disc-halo interface

Alexander Hobbs<sup>1</sup> & Robert Feldmann<sup>1</sup>

<sup>1</sup>*Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*

20 January 2020

## ABSTRACT

The flat star formation (SF) history of the Milky Way (MW) requires gas in the Galactic disc to be replenished, most likely from a reservoir outside the Galaxy. Such a replenishment may be achieved by a form of ‘positive’ feedback, whereby SF feedback creates a Galactic fountain cycle that collects and cools additional gas from the hot halo surrounding the Galaxy. In this paper we present a model of this process for the MW. A section of the Galactic disc is allowed to form stars which subsequently explode as supernovae (SNe) and send gas out into the hot halo. The gas that is sent out is colder than the hot halo gas and, as it mixes, the halo gas is cooled, providing fuel for further SF as the mixture falls back onto the Galactic disc. We find that this process can be sufficient to maintain a roughly-constant cold gas mass in the MW over at least 3 Gyr. Our results further suggest that there is a positive feedback trend whereby increasing SF leads to an increase in the cold gas budget at average SF rates below  $0.5 \text{ M}_\odot \text{yr}^{-1}$  but above which becomes negative, where further increasing the SFR causes the cold gas budget to decrease. We have constructed an analytical model for this that reproduces the data well and could have profound implications for galaxy evolution in feedback-dominated regimes.

## Key words:

## 1 INTRODUCTION

The presence of a hot, gaseous halo around the Milky Way (MW) and other, late-type spiral galaxies (for the lines of evidence for this see, e.g., Kacprzak et al., 2008Gaensler et al., 2008Nicastro et al., 2008Mastropietro et al., 2005Grevich & Putman, 2009Bogdán et al., 2013Miller & Bregman, 2015Anderson et al., 2016) has opened up a realm of scientific inquiry into the relationship between the galactic disc and the hot, low-density gas that surrounds it. A link between this halo gas and the star-forming gas in the galaxy exists through a process known as a ‘galactic fountain’, whereby feedback from supernovae (SNe) in the disc causes gas to be ejected out of the plane and into the disc-halo interface which extends to  $\sim 2 - 3$  kpc. The ejected gas may later fall back onto the disc due to gravity (Shapiro & Field, 1976Bregman, 1980Melioli et al., 2008) and provide fresh fuel for star formation, allowing the cycle to repeat again.

This cycle of ejection and re-accretion has been touted as a possible explanation for the near-constant SF rate (SFR) in our own Galaxy over the last  $\sim 8$  Gyr since  $z \sim 1$  (e.g., Rocha-Pinto et al., 2000). This statistic stands in contrast to the SFR of the wider Universe, which shows a dramatic decline since redshift  $z \sim 2 - 3$  (e.g., Lilly

et al., 1996Madau et al., 1998Hippelein et al., 2003). A near-constant rate implies that the gas used up in star formation (SF) must be constantly replenished, over nearly a Hubble time. Accretion from the larger cosmological environment can likely only play a part until  $z \simeq 1$  (Kereš et al., 2005Fernández et al., 2012Nelson et al., 2013Woods et al., 2014), and mergers with other galaxies likely only account for a  $\sim$  tenth of what is required (see Sancisi et al., 2008, for a review). Recycling of gas by stellar winds may be able to account for closer to half of what is required at low redshift ( $z < 0.5$ ) but it is not clear whether this gas would be returned to the disc or would end up in the halo (Leitner & Kravtsov, 2011). Wind recycling moreover suffers from a law of diminishing returns and would be unable to provide the required low-metallicity gas to remedy the G-dwarf problem (Schmidt, 1963Worthey et al., 1996) and related K- and M-dwarf problems (Casuso & Beckman, 2004Woolf & West, 2012). There must therefore be an alternative mechanism, not yet established in the scientific community, and it likely relates to the galactic fountain process – a form of ‘positive feedback’ whereby the action of supernovae (SNe), rather than to get rid of gas, is to bring more gas into the Galaxy that can contribute to the star-forming gas budget.

Observations of the disc-halo interface have begun to provide a wealth of information in recent years. Surveys have unmasked the presence of extra-planar cold atomic (H I) gas

\* E-mail: alexander.p.hobbs@gmail.com

in the vicinity of the MW (for a review see, e.g., Putman et al., 2012). Of particular interest are those extra-planar structures termed ‘intermediate-velocity clouds’ (IVCs), that possess radial velocities with respect to the local standard of rest (LSR) of  $40 \text{ km s}^{-1} \leq |v_{\text{LSR}}| \leq 90 \text{ km s}^{-1}$ . The IVCs inhabit a region in the halo within  $\sim 2 \text{ kpc}$  from the disc plane (Röhser et al., 2016). Their properties, including their near-solar metallicity, suggest that they may be connected to a fountain process occurring in our own Galaxy. A class of IVCs of particular interest are those that contain significant amounts of molecular gas, allowing them to be detected in carbon monoxide (CO) emission (Magnani & Smith, 2010). These are known as molecular IVCs (MIVCs), and they appear to be quite numerous. Röhser et al. (2016) employ a method using far-infrared (FIR) excess emission to probe small angular scales, and find 239 MIVC candidates over both Galactic hemispheres with both negative and positive radial velocities, within a latitude limit of  $20^\circ$ . This distribution of MIVC candidates is in good agreement with the rotational model of Marasco & Fraternali (2011), which already lends support to the picture of IVCs as originating from galactic fountain gas. Extrapolating to the entire MW, Röhser et al. (2016) find that the infalling IVC and MIVC gas may account for the main fraction of the gas inflow onto the disc. MIVCs may therefore be a late stage in the Galactic fountain process.

Recently, the presence of molecular gas in galaxies has been found to correlate with the mass, and density, of the hot ‘atmosphere’ surrounding giant ellipticals and early spiral galaxies (Babyk et al., 2018). The current interpretation of this finding is that cooling from the hot atmosphere is the origin of the molecular gas. In particular, it is a cooling time of  $< 1 \text{ Gyr}$  within  $10 \text{ kpc}$  that allows for the formation of molecular gas; this is the ‘cooling time threshold’ (Cavagnolo et al., 2008; Rafferty et al., 2008; Hogan et al., 2017) which has been observed to be a reasonably sharp cut-off in the presence of molecular gas in early-types and brightest cluster galaxies (BCGs). In early-types, it is only possible to get molecular gas if you are below this cooling time threshold – but in spiral galaxies molecular gas is present regardless (e.g., Wilson, 2013). This points to the conclusion that in early-types it is the simple condensation of the hot corona that gives rise to molecular gas, but in spiral galaxies such as the MW there is an additional mechanism occurring that does the job. Recent observations of the massive outflow in NGC6240 have begun to probe the origins of the molecular gas, finding that a significant fraction of it is entrained in the outflow itself rather than residing in the galaxy (Cicone et al., 2018).

On the theoretical side, simulations have typically seen stellar feedback in the galactic disc as a form of negative feedback, by both heating gas and by ejecting gas out of the galaxy, thereby leaving less ‘cold’ gas available to form stars. Indeed, in low-mass halos, stellar feedback is invoked as resolving the discrepancy between the halo mass function and the galaxy mass function (e.g., Puchwein & Springel, 2013). However, for halos more massive than a few times  $10^{10} \text{ M}_\odot$  the picture of stellar feedback as playing a negative role in the gas supply of the galaxy has recently been challenged (Fraternali & Binney, 2008; Marinacci et al., 2011; Hobbs et al., 2013; Hobbs et al., 2015; Armillotta et al., 2016).

The problem with invoking a hot halo as the source of

the continued SF is that such a feature is extremely slow to cool and stable to thermal instabilities (Joung et al., 2012). Early simulations attempting to model galaxies with hot halos found erroneous cooling, usually taking the form of a sudden condensation into cold gas clumps, that is now known to be due to numerical instabilities in the classic smoothed particle hydrodynamics (SPH) technique (Hobbs et al., 2013; Nelson et al., 2013) and does not feature in simulations conducted with newer techniques that do not suffer from these (particular) numerical errors. Of course, without cooling, the gas is hydrostatically supported and cannot add to the cold gas budget of the disc at the required rate. If gas is to cool from the hot halo, therefore, some mechanism is needed that can excite a thermal instability. One such mechanism was proposed by Marinacci et al. (2011), in which a galactic fountain seeds metal-rich gas into the metal-poor hot halos, giving rise to a thermal instability that causes cold gas to rain down onto the disc in the form of  $\sim 10^5 \text{ M}_\odot$  clouds (see also Fraternali & Binney, 2008). This model provides an excellent fit to both the velocity and spatial distributions of warm IVC-like HI gas in the MW (Marasco et al., 2012). 2D and 3D simulations of individual cold ( $T \approx 10^4 \text{ K}$ ) clouds moving through a hot medium were presented by Marinacci et al. (2011) and subsequent follow-up papers (e.g., Armillotta et al., 2016), where the turbulent mixing of the cloud with the hot gas and subsequent cooling caused the mass of cold gas to increase by  $\sim 30\%$  over a period of 60 Myr. Similar work by Gronke & Oh (2018), investigating the formation of an extended ‘cooling tail’ that continues to draw in and cool hot gas through thermal pressure gradients, found an increase in dense gas mass in the cloud of an order of magnitude over  $\approx 20$  cloud-crushing times.

Such clouds are presumed to be the ejecta from the disc via SNe feedback, and travel through the hot halo, entraining and cooling halo gas before returning to the disc and delivering this gas as fuel for further SF. The key concept here is that the cooling of gas from the halo is therefore dependent on the presence of star-forming gas deeper inside the potential well. The Marinacci et al. (2011) Armillotta et al. (2016) papers argue that this model is sensitive to galactic environment, i.e., in halos above  $\sim 10^{13} \text{ M}_\odot$  the process is ineffective. This implies a quenching of SF and an inability to cool the hot corona going from late-type to early-type systems.

Such a mechanism constitutes a positive feedback mode of gas accretion onto the disc – where, rather than hindering SF and reducing the cold gas budget, feedback from SNe can do the opposite, acting as a means to ‘refrigerate’ the hot halo and encourage it to cool and accrete onto the galactic disc.

Many simulations of late-time star formation in galaxies – particularly those using early versions of SPH – were plagued by numerical artifacts that prevented a clear picture of how galactic cold gas accretion proceeds. There has been much discussion, debate, and development in this area in the past decade, however, and the latest codes are substantially more reliable (Hopkins, 2015), allowing for useful experimentation to be carried out with galaxy formation models. Such models now possess improvements in physical modelling techniques, such as second-order accurate radiative cooling, chemical evolution, improved SF and feedback with metallicity evolution, and magnetic fields. The code we

employ is arguably one of the most up-to-date and competitive codes in this regard, particularly in relation to treating fluid instabilities and mixing processes in astrophysical gases.

Our work stands apart from the papers referenced above in that it simulates a Galactic setup in three dimensions, rather than following an individual cloud or selection of clouddlets. The outflow is generated self-consistently from SF and subsequent SNe explosions within the Galactic disc, and we follow the Galactic fountain cycle over a period of Gyrs.

## 2 METHOD

Our simulations are run with our own modified version of the public GIZMO code (Hopkins, 2017b). GIZMO is a mesh-free hybrid between Lagrangian and Eulerian solvers, designed to benefit from the advantages inherent to each. It discretizes the equations of hydrodynamics on a set of discrete tracers which represent unstructured cells, and assigns the volume of the cells based on a kernel function. The flux between the cells is determined by solving the Riemann problem across the boundary, avoiding the need for artificial dissipation terms.

The hybrid mesh-free approach taken by GIZMO allows for the mixing of different fluid phases naturally (as the flux of all fluid quantities is calculated between cells as default), while at the same time retaining the ability for Lagrangian adaptivity of cells and spatial resolution, and maintaining exact mass, energy, and momentum conservation. The meshless-finite-mass (MFM) approach that we use is also second-order consistent, which allows for formal convergence without needing to tend to the limit of an infinite neighbour number.

Radiative cooling in GIZMO is performed down to 10 K using the cooling tables of Katz et al. (1996). The cooling prescription includes a uniform extragalactic ultra-violet (UV) background (Haardt & Madau, 2012). We include full physical thermal conduction with a fiducial conduction coefficient of 0.1 of the Spitzer value (Spitzer, 1962), and physical viscosity set to the Braginskii viscosity (Hopkins, 2017c). We also employ turbulent diffusion of metals across cells as per the Smagorinsky turbulent eddy diffusion model (Hopkins, 2017a; Colbrook et al., 2017). The gas has self-gravity in addition to an external analytic potential (see next Section).

### 2.1 Initial conditions

Our model consists of a section of the MW disc, embedded in a hot ambient medium of low-density halo gas. The disc section has constant surface density in  $x$  and  $y$ , and a Gaussian profile in the vertical direction  $z$ . We use periodic boundary conditions in the  $xy$  plane, with an open boundary in the vertical ( $z$ ) direction, and position the disc plane at the centre of the box, so that the halo gas is simulated both above and below the plane. This allows the model to cater for (i) differences in the outflow that is generated above and below the plane, and (ii) the effect on the outflow gas if any passes back through the disc and out the other side. A visualisation of the setup at an early time is shown in the left-most panel of Figure 1.

The setup was chosen based on the model in the SILCC

papers (Walch et al., 2015), which approximates a present-day MW disc, although our implementation is 16 times larger in the disc plane in order to explore a larger range of behaviour. The initial conditions (ICs) comprise a periodic box of gas measuring  $2 \text{ kpc} \times 2 \text{ kpc} \times \pm 50 \text{ kpc}$  (the extreme length in the  $z$ -direction is to ensure that any numerical effects at the open boundary affect the behaviour of the gas near the disc as little as possible). The disc has a midplane density of  $9 \times 10^{-24} \text{ g cm}^{-3}$  and a scaleheight of 60 pc. The disc vertical extent is cut off at the height where the disc density becomes equal to the ambient density,  $|z| \approx 0.24 \text{ kpc}$ . The disc gas has solar metallicity. The total mass of the disc gas is  $4 \times 10^7 \text{ M}_\odot$ . The halo gas has a density of  $0.0006 \text{ atoms cm}^{-3}$  and a metallicity of 0.1 solar. In addition to the gas disc, a stellar disc analytic potential is included which has the form of a  $\text{sech}^2$  profile – the ‘isothermal sheet’ (Spitzer, 1942), which is described by equation 29 in Walch et al. (2015). The stellar disc potential has a surface density of  $30 \text{ M}_\odot \text{ pc}^{-2}$  and a scaleheight of 100 pc.

Our IC is set up to be initially in hydrostatic equilibrium with both the gas distribution and the external analytic gravitational potential. The temperature in the disc varies in the vertical direction,  $T \equiv T(z)$ , from  $\sim 10^1 - 10^5 \text{ K}$  (the majority being at  $\approx 5 \times 10^2 \text{ K}$ ). The temperature in the halo (within 10 kpc) is  $T \approx 10^6 \text{ K}$ .

For most of the ICs (including the fiducial) we seed a one-off, initial turbulent velocity field in the disc that is constructed based on the treatment in Dubinski et al. (1995). Briefly, the velocity field follows a Kolmogorov power spectrum,

$$P_v(k) \sim k^{-11/3}, \quad (1)$$

where  $k$  is the wavenumber. Such a velocity field is homogeneous and incompressible, and so we can define  $\mathbf{v}$  in terms of a vector potential  $\mathbf{A}$  such that  $\mathbf{v} = \nabla \times \mathbf{A}$ . The components of  $\mathbf{A}$  can then be described by a Gaussian random field with an associated power spectrum,

$$P_A(k) \sim k^{-17/3}. \quad (2)$$

The variance in  $|\mathbf{A}|$  diverges sharply as  $k$  decreases, and so we introduce a small-scale cut-off  $k_{\min}$ . We can then write:

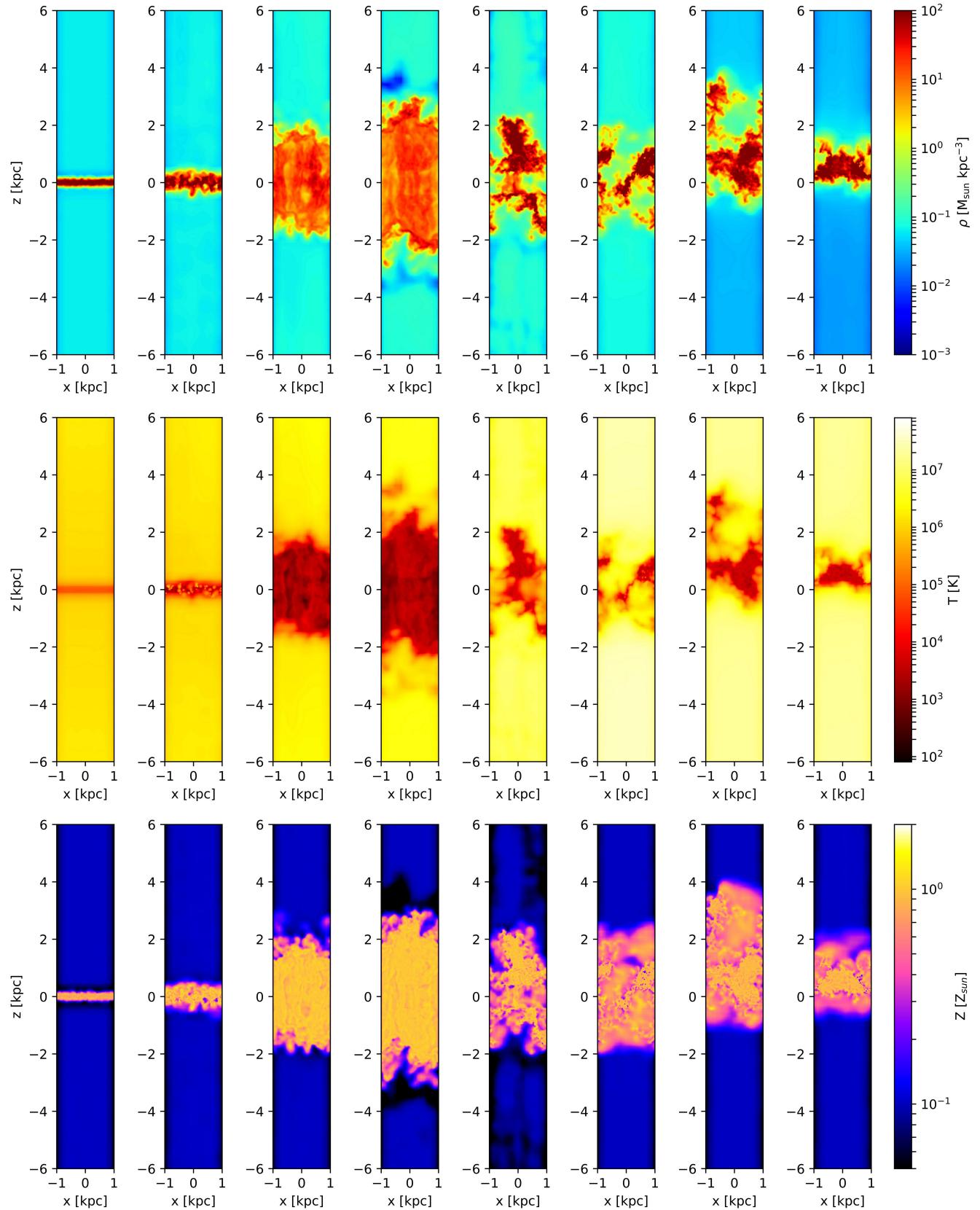
$$\langle |\mathbf{A}_k|^2 \rangle = C(k^2 + k_{\min})^{-17/6}, \quad (3)$$

where  $C$  is a constant that sets the normalization of the velocities. We set this equal to unity and normalize the velocity field once the statistical realization has been generated. Physically, the small-scale cut-off  $k_{\min}$  can be interpreted as the scale  $R_{\max} \simeq k_{\min}^{-1}$ , the largest scale on which the turbulence is likely to be driven. In our model we use  $R_{\max} = 1 \text{ kpc}$ .

To generate the statistical realization of the velocity field we sample the vector potential  $\mathbf{A}$  in Fourier space, drawing the amplitudes of the components of  $\mathbf{A}_k$  at each point  $(k_x, k_y, k_z)$  from a Rayleigh distribution with a variance given by  $|\mathbf{A}_k|^2$  and assigning phase angles that are uniformly distributed between 0 and  $2\pi$ . We then take the curl of  $\mathbf{A}_k$ ,

$$\mathbf{v}_k = i\mathbf{k} \times \mathbf{A}_k, \quad (4)$$

to obtain the Fourier components of the velocity field. Finally we take the Fourier transform of this to obtain the velocity field in real space. We use a periodic cubic grid of



**Figure 1.** Visualisation for our fiducial SF10-turb run, showing density (top), temperature (middle), and metallicity (bottom) in the side-on view at 8 different times:  $t = 1, 3, 10, 100, 200, 300, 400$ , and  $500$  Myr, left-to-right.

**Table 1.** Summary of the simulations in this paper.  $N_{\text{gas}}$  is the number of gas particles,  $m_{\text{gas}}$  is the (constant) mass of each gas particle,  $\eta_{\text{Sp}}$  is the thermal conductivity in units of the Spitzer conductivity,  $\nu_{\text{Brag}}$  is the physical viscosity in units of the Braginskii viscosity, and  $\rho_{\text{SF}}$  is the threshold density for star formation in atoms per cc.

Simulation ID	$N_{\text{gas}}$	$m_{\text{gas}} (\text{M}_\odot)$	$\eta (\eta_{\text{Sp}})$	$\nu (\nu_{\text{Brag}})$	$\rho_{\text{SF}}$ (atoms per cc)	Turbulent IC
H1	$2.2 \times 10^6$	23	0.1	1.0	10.0	Yes
H1-sf5	$2.2 \times 10^6$	23	0.1	1.0	5.0	Yes
H1-sf1	$2.2 \times 10^6$	23	0.1	1.0	1.0	Yes
H1-sf0.1	$2.2 \times 10^6$	23	0.1	1.0	0.1	Yes
H1-sf0.05	$2.2 \times 10^6$	23	0.1	1.0	0.05	Yes
H1-sf0.02	$2.2 \times 10^6$	23	0.1	1.0	0.02	Yes
H1-sf0.01	$2.2 \times 10^6$	23	0.1	1.0	0.01	Yes
H1- $\eta$ 1.0	$2.2 \times 10^6$	23	1.0	1.0	10.0	Yes
H1- $\eta$ 0.01	$2.2 \times 10^6$	23	0.01	1.0	10.0	Yes
H1- $\nu$ 0.1	$2.2 \times 10^6$	23	0.1	0.1	10.0	Yes
H1- $\nu$ 0.01	$2.2 \times 10^6$	23	0.1	0.01	10.0	Yes
M1	$1.3 \times 10^6$	38	0.1	1.0	10.0	Yes
L1	$6.6 \times 10^5$	77	0.1	1.0	10.0	Yes
H2	$2.2 \times 10^6$	23	0.1	1.0	10.0	No

dimension  $256^3$  when generating the statistical realization of the velocity field and we use tri-cubic interpolation to estimate the components of the velocity field at the position of each particle/cell.

Although the setup of our model is based on the SILCC model (Walch et al., 2015), our focus and science goals differ greatly. The SILCC papers deal with the state of the Galactic interstellar medium (ISM) for the purposes of understanding SF within the disc; we are interested in the mixing of disc gas and halo gas far from the disc in order to understand positive feedback. Our implementation also stands apart by employing a meshless hydro solver.

With the small section of the MW disc that we model, we are able to achieve a considerably higher resolution than if we modelled the entire Galaxy – in our fiducial runs, this is  $23 \text{ M}_\odot$  per cell. The resolution of all the runs is shown in Table 1.

## 2.2 Star formation and feedback

Star formation in our simulations proceeds subject to a density threshold that differs by 3 orders of magnitude ( $0.01 - 10$  atoms  $\text{cm}^{-3}$ ) between different runs and the molecular hydrogen fraction criterion of Hopkins (2017a) Krumholz & Gnedin (2011). The latter criterion is particularly useful at SF density thresholds  $< 100$  atoms  $\text{cm}^{-3}$  in correcting the lack of self-shielding molecules at low resolution. Feedback from stars is modelled according to Kim et al. (2014), and injects thermal energy and metals from an IMF-averaged number of type II SNe in a single burst after a main sequence time. The injected energy is weighted over the kernel. We note that this is a relatively simple SF and feedback model, but it has been shown to be effective in high reso-

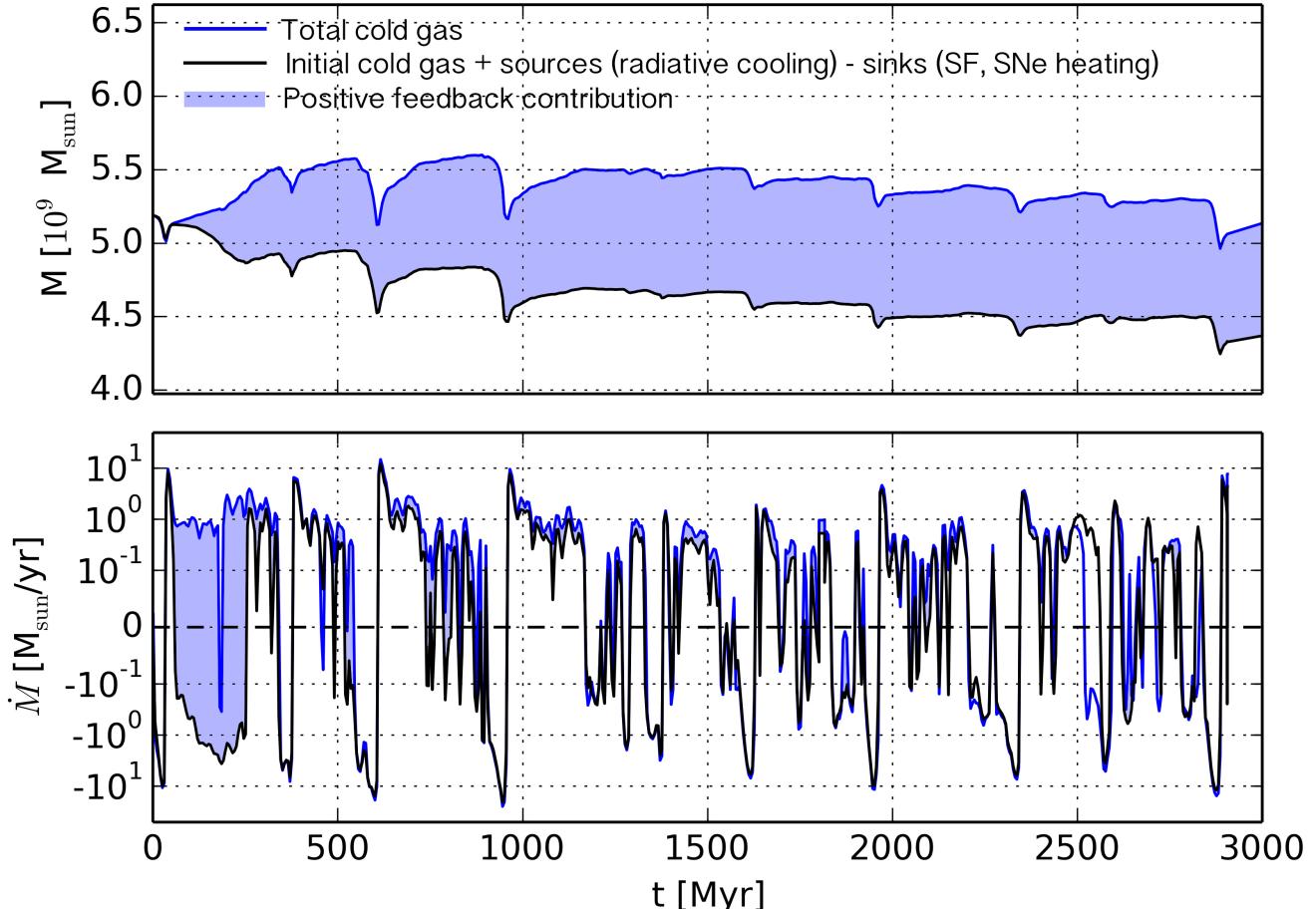
lution simulations with particle masses  $< 100 \text{ M}_\odot$  (Hopkins et al., 2018). As listed in Table 1, the majority of our simulations have a mass resolution of  $23 \text{ M}_\odot$ , are therefore able to make effective use of this model.

## 3 RESULTS

In Figure 1 we show a time-series with the state of the fiducial (H1) model at 8 different times:  $t = 1$  Myr,  $t = 3$  Myr,  $t = 10$  Myr,  $t = 100$  Myr,  $t = 200$  Myr,  $t = 300$  Myr,  $t = 400$  Myr, and  $t = 500$  Myr. In this Figure we show the density (top), temperature (middle) and metallicity (bottom), all averaged along the line-of-sight.

The initial star formation and subsequent SNe explosions taking place in the disc give rise to a lifting of some of the disc gas out to  $\sim 2 - 3$  kpc. The material is not ejected, however, since the outflow velocity is not high enough to overcome both the gravitational potential and the confining pressure of the hot gas. This results in the previously-lifted gas falling back and re-forming the disc, at which point it undergoes a subsequent starburst and the cycle repeats. This cycle repeats approx. every 300 Myr. The maximum height reached of the disc gas (the outflow) varies across the different runs, and is in particular sensitive to the strength of the initial star formation.

Each time the disk reforms, its structure and morphology are different, which naturally causes differences in the morphology of the SF and subsequent feedback. These differences compound as the simulation progresses, and so we note that these simulations are firmly in a somewhat chaotic regime whereby small changes between setups can amplify and give rise to large differences at later times. As a result we do not investigate morphologies in any detail, or make



**Figure 2.** Total amount of cold gas mass in the entire simulation over 3 Gyr (top) and the corresponding time derivative (bottom). The top panel compares the total cold gas (blue line) with the amount of cold gas that there would be assuming no contribution from the ‘halo gas extraction’ positive feedback process – namely, the black line is the original amount of cold gas that is gradually used up by SF and heated by SNe, which also includes the (very small) increase from radiative cooling of the initial hot ambient gas. The grey shaded area therefore represents the contribution to the cold gas budget made by positive feedback.

comparisons between runs at specific times; rather, we look at global, averaged quantities where possible.

The initial condition was varied between initially turbulent and initially not-turbulent, with very little difference found in the behaviour of the gas over the run (see Section 3.4.1). The one-off seeded turbulence, while playing a role in creating some structure and density variation in the disc, is ultimately less effective than the SNe themselves in determining the dynamics of the outflow.

### 3.1 Cold gas: mass increase and distribution

**Note:** For all quantitative plots and recorded values we extrapolate from our  $2 \text{ kpc} \times 2 \text{ kpc}$  section of the MW disc to the entire disc – therefore, numbers quoted for the amount of cold gas, SF, the corresponding time derivatives, and all relevant figures, have been multiplied by  $M_{\text{disc}}/M_{\text{patch}}$ , where  $M_{\text{disc}}$  is the total mass of the present-day MW gaseous disc, and  $M_{\text{patch}}$  is the total mass of the section that we model. The conversion factor is  $M_{\text{disc}}/M_{\text{patch}} = 132.5$ . The division by  $M_{\text{patch}}$  gives us each relevant quantity per unit mass (in other words, a ‘specific’ rate for each quantity), which we

then scale up to the whole MW disc. This simplistic scaling is not meant as a substitute for a full scale modelling of a MW disk but rather as a means to facilitate a comparison of our results with observations. For instance, this scaling will clearly overpredict the time variation of the global SFR given that in reality different patches may experience a burst at different times.

The nature of positive feedback from SF is that the action of SNe actually gives rise to an *increase* in cold gas, rather than a decrease. As such, we track the amount of cold gas in the simulation over time, where ‘cold gas’ is defined as having  $T < 2 \times 10^4 \text{ K}$  (as per Armillotta et al., 2016). This is shown in Figure 2 for our H<sub>2</sub> run over 3 Gyrs (blue line). The amount of cold gas in the simulation is largely constant over the course of the 3 Gyrs, even with vigorous SF converting cold gas to stars and SNe heating the surrounding gas to high T. The SF in the simulation over the 3 Gyrs is shown in Figure 3. From the bottom panel of this Figure, which shows the SFR, one can see that the SF is bursty, alternating between periods of high SFR (reaching values of  $\gtrsim 1 M_{\odot} \text{ yr}^{-1}$ ) before shutting off and waiting for the disc to re-form.

The near-constant cold gas mass is therefore indicative of a ‘self-sustaining’ MW + hot halo ecosystem, whereby cold gas sent into the hot halo gas is able to mix and entrain the hot gas and cool it, adding to the reservoir of fuel for SF the next time the cycle repeats.

Of course, in such a model it can be difficult to separate out the relevant processes from each other and determine the contribution that each makes. To show the contribution from positive feedback alone, we have plotted in Figure 2, alongside the total cold gas mass, the amount of cold gas that would have been seen in the simulation should this positive feedback process not have added to the cold gas budget. This is the black line in Figure 2, and it was calculated by taking the initial cold gas mass and (i) subtracting the SF and SNe heating over the lifetime of the simulation (the cold gas ‘sinks’) and (ii) adding the radiative cooling of the hot halo gas (a cold gas ‘source’). The difference between the black and the blue lines on this plot (the shaded blue area) is therefore the contribution from positive feedback alone. It is clear that the vertical height of this shaded area grows quickly in the first 400 Myr, after which it remains largely the same. We discuss this initial vs. latter phase and the reasons for it in Section 4.3.

### 3.2 Positive vs. negative feedback

To further investigate this process of positive feedback from SF, and in particular, to compare it to the usual negative feedback that is caused by the heating of the gas through SNe, we tracked the hot halo gas and the disc gas separately through their IDs. We define a metric for positive feedback as gas that was originally in the hot halo phase achieving a temperature below  $2 \times 10^4$  K (temperature threshold defined as per Armillotta et al., 2016). We also define a metric for negative feedback; gas that was originally in the disc reaching a temperature above  $2 \times 10^5$  K (this was originally the highest temperature of the gas that made up our Gaussian disc IC). In this way, we can measure how much of the initially-hot gas became cold, and how much of the initially-cold gas became hot. Since the cooling time for gas  $> 10^6$  K is long, we can be confident that most of the cooled hot halo gas is the result of the disc gas mixing with the halo gas (and indeed, we have verified this by running a stand-alone model with only the hot ambient gas allowed to cool radiatively with no influences from the disc). The initially-cold gas in the disc that gets heated is the result of the SNe injecting thermal energy into their surroundings.

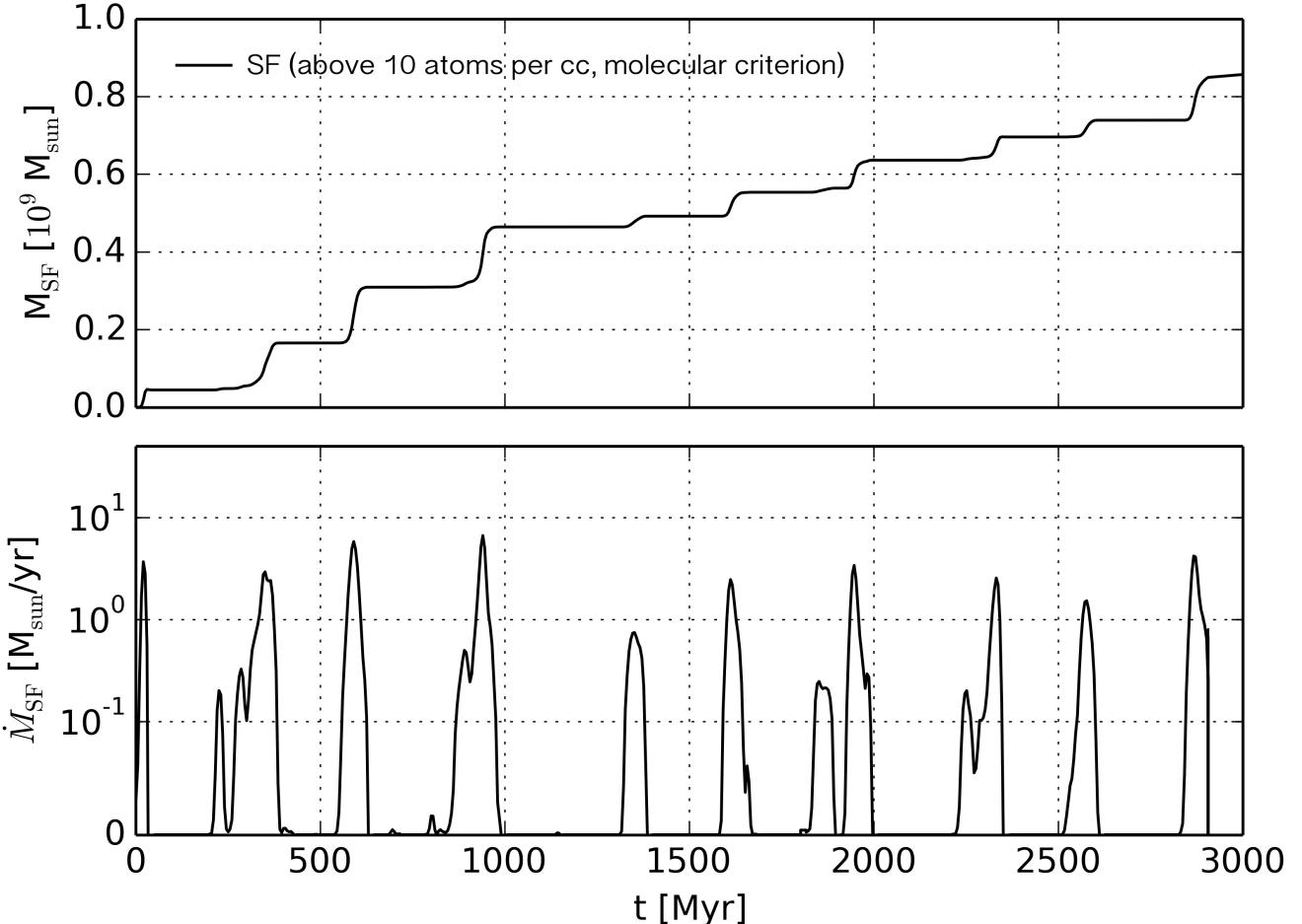
Figure 4 shows these two metrics over 1 Gyr for our H1 (fiducial) run. The top panel shows the total mass of ‘extracted’ gas, and the bottom panel shows the time derivative. The initially-disc gas that is heated up is naturally somewhat bursty, since the SNe heating coincides with the semi-regular starbursts (see, e.g., Stinson et al., 2007). The cold gas extracted from the hot halo (via positive feedback) grows faster than hot gas created from the disc, showing that positive feedback metric exceeds negative feedback in this setup. The rate-of-change of the cold/hot gas extracted from the halo/disc differs quite clearly, with the extracted cold gas rate remaining mostly above zero (with occasional dips) and the extracted hot gas oscillating above and below zero far more often. The peak values of  $\dot{M}_{\text{extr}}$ , however, are similar.

### 3.3 Varying the SF threshold

Since both the positive feedback and the negative feedback mechanism depend on SF, it makes sense to explore varying SF rates. While the SF rate is a quantity that arises naturally through the behaviour of the cold gas in the simulation, and cannot be set directly by hand, we can use the SF threshold (listed in Table 1 for each run) as a proxy. In varying the density threshold at which stars are allowed to form (although still subject to the molecular gas fraction criterion) we can achieve different average SF rates across the lifetime of the simulation. The different SF thresholds give rise to different amounts of cold gas extracted from the halo via the positive feedback process. These are shown in Figure 5. The different runs produce roughly the same amount of extracted cold gas (top panel) and the corresponding time derivative (bottom panel) up to  $\approx 200$  Myr, after which they begin to diverge. This point in time coincides roughly with the second starburst and the first repetition of the fountain cycle, and indeed small differences in the initial amount of star formation can lead to far larger differences each time the cycle is repeated. In general, the rates of cold gas extraction for each of the varying-SF threshold runs have a similar form post 200 Myr, but the locations and frequency of the dips below zero vary, with the lowest SF threshold runs having the most dips into negative  $\dot{M}_{\text{extr}}$ .

Figure 6 shows the mass of stars formed (top panel) and the SFR (bottom panel) for this suite of runs varying the SF threshold. A lower SF threshold naturally leads to a larger starburst at  $t = 0$ , although the peak value reached saturates at 0.1 atoms per cc. The major difference however is clearly seen in how bursty the SFR is across the different runs. At a high SF threshold (10 atoms per cc run in our fiducial run), SF is restricted to high density gas and, thus, can only take place in the gas disk. In contrast, at a low SF threshold, stars are able to form more easily in outflowing gas and at higher altitudes above the midplane. This has the effect of creating a more continuous SF history over the run, although the typical maximum SFR values reached are similar. However, the more sustained SFR for the thresholds at the lower end of range result in a total stellar mass that is  $\approx 4$  times higher than our fiducial 10 atoms per cc run. In general, a higher SF threshold returns a lower average SF rate across 1 Gyr, as seen in table 2. The relationship is not proportional, however, as the behaviour of the gas in the simulation is extremely chaotic and influenced by many different factors.

From Figure 5 we see that, above  $\rho_{\text{SF}} = 1$  atom cm $^{-3}$ , the amount of extracted cold gas decreases as the SF threshold is increased. Below  $\rho_{\text{SF}} = 1$  atom cm $^{-3}$ , this trend is reversed, with a higher SF threshold leading to *more* cold gas extracted from the halo. This is an interesting result, and, given the inverse scaling between SFR and SF threshold, it leads us to offer an alternative definition of positive feedback; that of a positive feedback **trend**, whereby an increase in SFR gives rise to an increase in the cold gas extraction rate. Similarly, we define a ‘negative feedback trend’ as one in which an increase in SFR in the Galaxy leads to a decrease in the rate of cold gas extraction (as is the usual interpretation of the effect of SNe on the Galactic cold gas budget). The bottom panel of Figure 5 shows the rate of cold gas mass extracted from the hot halo, and from this we

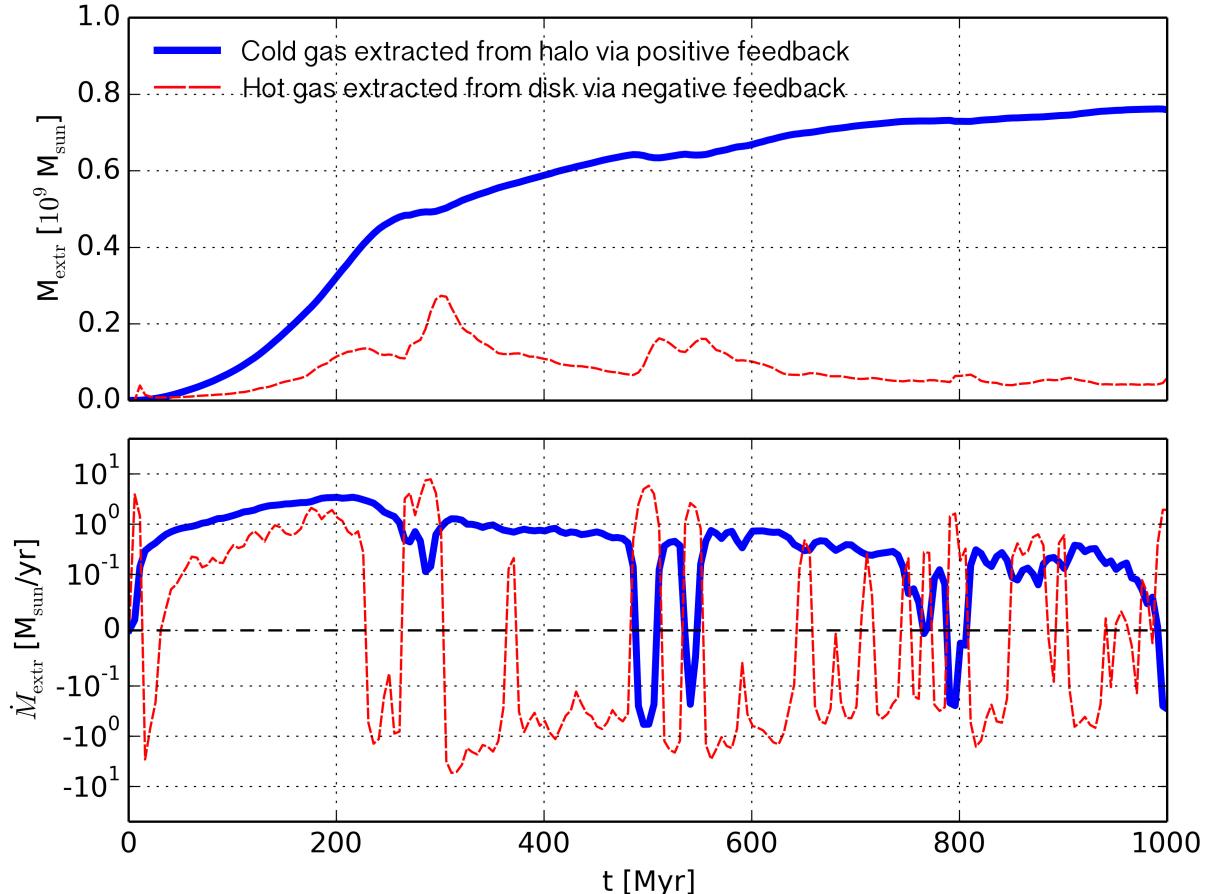


**Figure 3.** Total amount of mass in newly-formed stars in the entire simulation over 3 Gyr (top) and the corresponding rate (bottom). Here it is clear that we have a Galactic fountain cycle repeating many times, with each peak in the SFR (bottom panel) corresponding to the disc re-forming before a subsequent starburst that blows the gas out for  $\approx 200 - 300$  Myrs, during which time there is very little to no SF.

can get an average rate over the run. These values are listed in Table 2 for each value of the SF threshold. In Figure 6 we see the mass of SF (top panel) and the corresponding rate (bottom panel) for the same set of runs. We can extract an average SF rate for each run from this Figure, and these are also listed in Table 2.

Nonetheless, there is a general trend of lower average SFR with higher SF threshold. We therefore have 7 different values of the average SFR over 1 Gyr, which we can plot against the average rate of cold gas extraction over the same period. This is shown in Figure 7. When we do this, we find that at SFRs below  $\approx 0.5 M_{\odot} \text{ yr}^{-1}$ , the extracted cold gas rate and thus the cold gas budget for the Galaxy increases with increasing SFR. This demonstrates in a simple plot our 'positive feedback trend' below  $0.5 M_{\odot} \text{ yr}^{-1}$ . Above this SFR value, we see that the behaviour of the extracted cold gas rate with SFR follows a negative feedback trend – increasing the SFR above  $0.5 M_{\odot} \text{ yr}^{-1}$  extracts less and less cold gas from the hot halo. This Figure will be discussed in detail in Section 4, as it has interesting implications for galaxy evolution if we are to assume that positive feedback processes play a significant role.

The next plot, Figure 8, shows the same two positive feedback/negative feedback regimes but with the average rates drawn from either  $t < 400$  Myr or  $t > 400$  Myr, since it is clear from all the previous plots vs. time that the behaviour of the outflow in terms of extracting cold gas is significantly different within the first 400 Myr compared to later times. As we remark on in Section 4, this is not due to an initial transient arising from the IC, but rather a consequence of the favourable conditions for cold gas extraction within the first 400 Myr of the simulations. In Figure 8 we see indeed that the positive feedback trend peaks at a higher value,  $\approx 2 M_{\odot} \text{ yr}^{-1}$ , for  $t < 400$  Myr than when averaged over the whole Gyr. The trend for  $t > 400$  Myr also shows a peak near SFR  $\sim 0.5 M_{\odot} \text{ yr}^{-1}$ , although the peak is much less pronounced and the cold gas extraction rate shows only a weak scaling with SFR for low SFRs. This explains why we see only weak effects of positive feedback after  $t = 400$  Myr.



**Figure 4.** Positive vs. negative feedback comparison for H1 (fiducial). Extracted gas mass (top panel) and the corresponding time derivative (bottom panel). The cold gas extracted from the hot halo is several times larger than the hot gas extracted from the disc for the majority of the run. The reason for this can be seen in the bottom panel, where although both the cold and hot extracted gas reach similar peaks and troughs, the extracted cold gas rate stays mostly in the upper (positive) half while the extracted hot gas rate oscillates almost equally above and below zero.

### 3.4 Varying the model: turbulence, physical parameters, and resolution

In this section we compare the results for the extracted cold gas mass and SF (and the corresponding time derivatives) over 1 Gyr while varying numerical aspects of the model to test for convergence.

#### 3.4.1 Varying the IC

The first is the setup of the initial condition – we have run two different ICs, one with a turbulent velocity field in the disc (see Section 2.1) and one without. This is important for determining to what extent the details of the IC play a role in how the outflow gas extracts and cools gas from the halo. We find that there is very little difference between the two setups when it comes to the mass of cold gas extracted from the ambient hot gas or the stellar mass as shown in Figure 9. The rates of each differ only in terms of the location of the peaks/troughs, due to the somewhat chaotic nature of

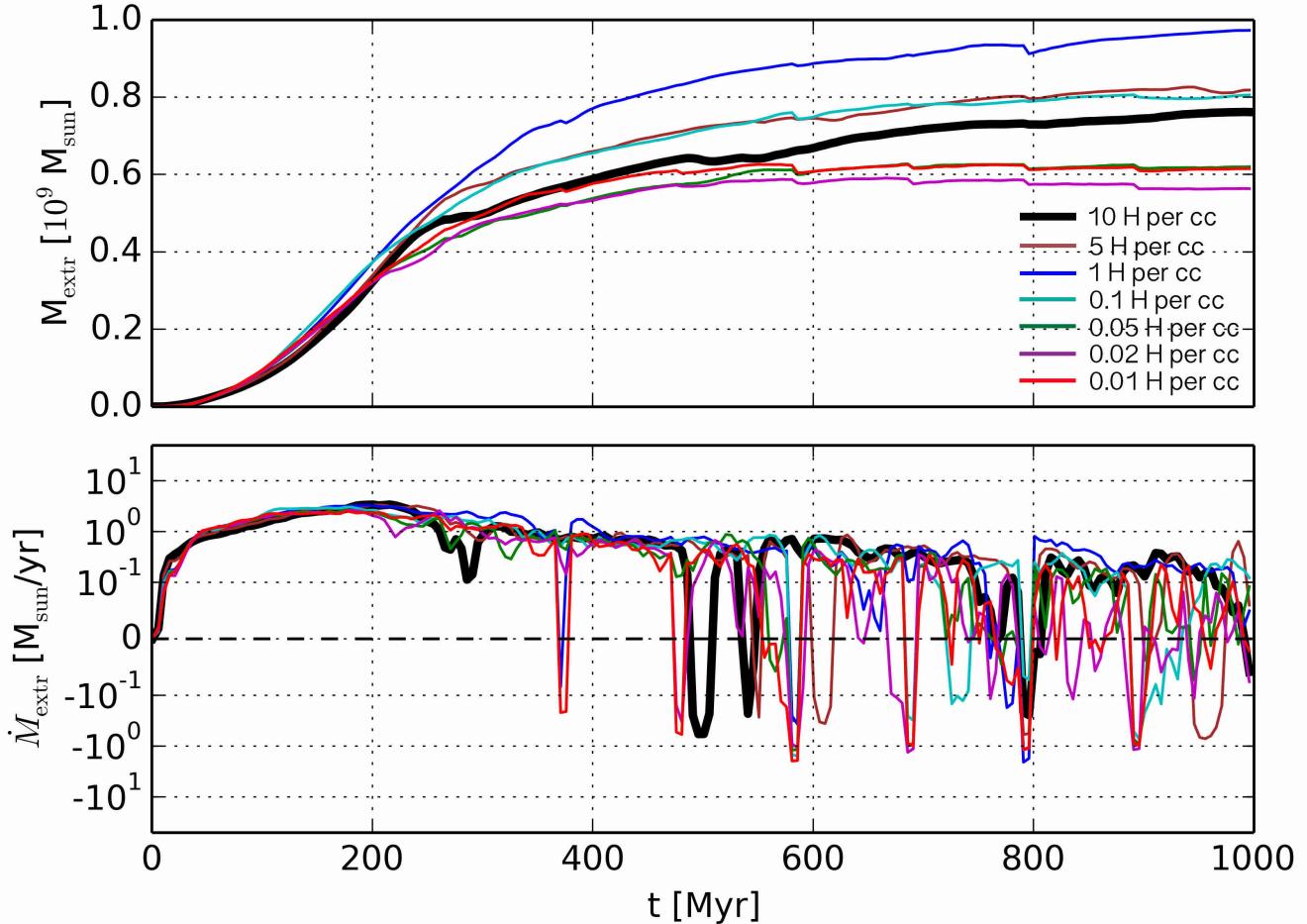
the simulation and the susceptibility of the starburst cycle to small changes in the parameters.

#### 3.4.2 Varying the resolution

Figure 10 shows the extracted cold gas and SF over 1 Gyr for 3 different resolutions, as listed in Table 1 (H1, M1, and L1 runs). These were all performed with our fiducial, turbulent IC. The two lower resolution simulations produce slightly more cold gas by the end of the run at 1 Gyr, but this is mostly due to a higher rate of cold gas increase for M1 and L1 between  $\approx 300 - 500$  Myr, after which the gradients are roughly the same across all three resolutions.

#### 3.4.3 Varying the thermal conductivity

We also explore the effect of varying the thermal conductivity parameter,  $\eta$ , which refers to the fraction of the Spitzer conductivity that particles in the simulation possess. Figure 11 shows three values of  $\eta$ , 1.0, 0.1, 0.01 (fiducial) and 0.01.



**Figure 5.** Plot showing the mass of extracted cold gas (top) and the corresponding time derivative (bottom), comparing simulations H1-sf with SF thresholds as listed in the key. The fiducial H1 run is shown with a thick black line for clarity. For the high SF thresholds, a reduction in the threshold leads to a greater amount of cold gas by the end of the run; however this trend reverses as the SF threshold drops below 1 atom cm $^{-3}$ . For the very low SF thresholds (0.01, 0.02 atoms cm $^{-3}$ ), the amount of extracted cold gas starts to slowly decrease again towards the end of the run.

There appears to be no obvious trend in how the thermal conductivity affects the amount of cold gas collected, or the mass of stars formed, and the total mass of extracted cold gas after 1 Gyr varies by  $\approx 2 \times 10^8 M_{\odot}$ , although the  $\eta = 0.1$  and  $\eta = 0.01$  runs are converged up to 400 Myr.

### 3.4.4 Varying the physical viscosity

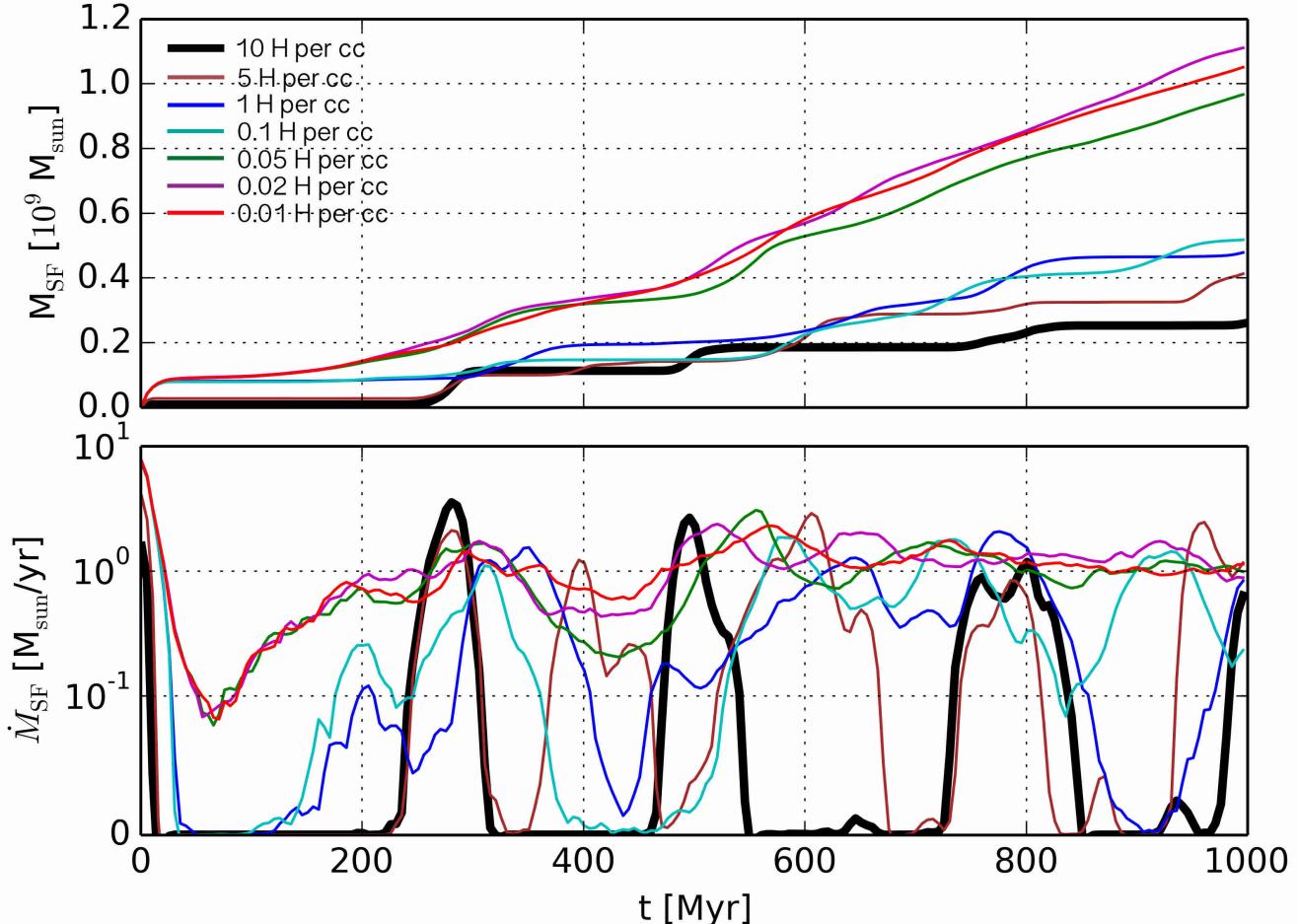
Finally, we vary the normalisation of the physical viscosity  $\nu$ , as a function of the Braginskii viscosity. Figure 12 shows three values of  $\nu$ , 1.0 (fiducial), 0.1, and 0.01. The physical viscosity parameter seems to have very little effect on the total amount of cold gas collected by 1 Gyr, with the top panel very well converged. The corresponding time derivatives differ slightly in where the dips are but on average are similar. The star formation is converged but naturally differs after  $t = 400$  Myr as to when the starbursts occur.

The purpose of this section was to compare the effect of varying a few of the parameters, but it should be noted that we kept the SF threshold fixed at the fiducial value of 10 atoms per cc. As such we are not investigating the SFR-

cold gas extraction trend that we discuss in Section 4.2 and display in Figures 7 & 8. Without running an exhaustive suite of simulations the effect of varying these parameters remains unknown – it is possible that both the position and the normalisation of the peak separating the positive and negative feedback trends would be affected.

## 4 DISCUSSION

The action of SNe feedback promoting additional cooling is a departure from the conventional wisdom on the role of SNe outflows in galaxy formation and evolution. Typically, galaxies have been modelled as some form of ‘bathtub’, with source and sink terms (Lilly et al., 1996Dekel & Mandelker, 2014) where SNe outflows exclusively remove cold gas from a galaxy. This constitutes ‘negative feedback’, which has been the accepted norm. Positive feedback, which is a somewhat more recent idea (Marinacci et al., 2011Frernali & Tomassetti, 2012Armillotta et al., 2016Hobbs et al., 2016) complicates this picture by introducing, via a galactic fountain pro-



**Figure 6.** Plot showing the mass of stars formed (top) and the SFR (bottom), comparing simulations H1-sf with SF thresholds as listed in the key. The fiducial H1 run is shown with a thick black line for clarity. In general, the high SF thresholds give bursty SF histories, while the low SF thresholds show a more continuous, high SFR over the course of the run.

cess operating in the hot gaseous halo, a means to increase the galaxy's store of cold gas available for star formation.

The mechanism by which this occurs is a mixing of cold gas initially located within the galactic disc with hot gas in the halo. The cold gas is sent into the hot halo gas by the action of SNe feedback, lifting it from the plane of the galaxy but not ejecting it from the galaxy's potential well, so that it later falls back after mixing with hot halo gas. This parcel of mixed hot and cold gas shows a net cooling, as the cooling rate at  $\sim 10^6$  K (the approx. value for the hot halo gas that interacts with the fountain gas) is approx. an order of magnitude lower than the cooling rate at  $10^5$  K. A parcel of gas loses energy  $E$  at a rate,

$$\frac{dE}{dt} = -n^2 \Lambda(T), \quad (5)$$

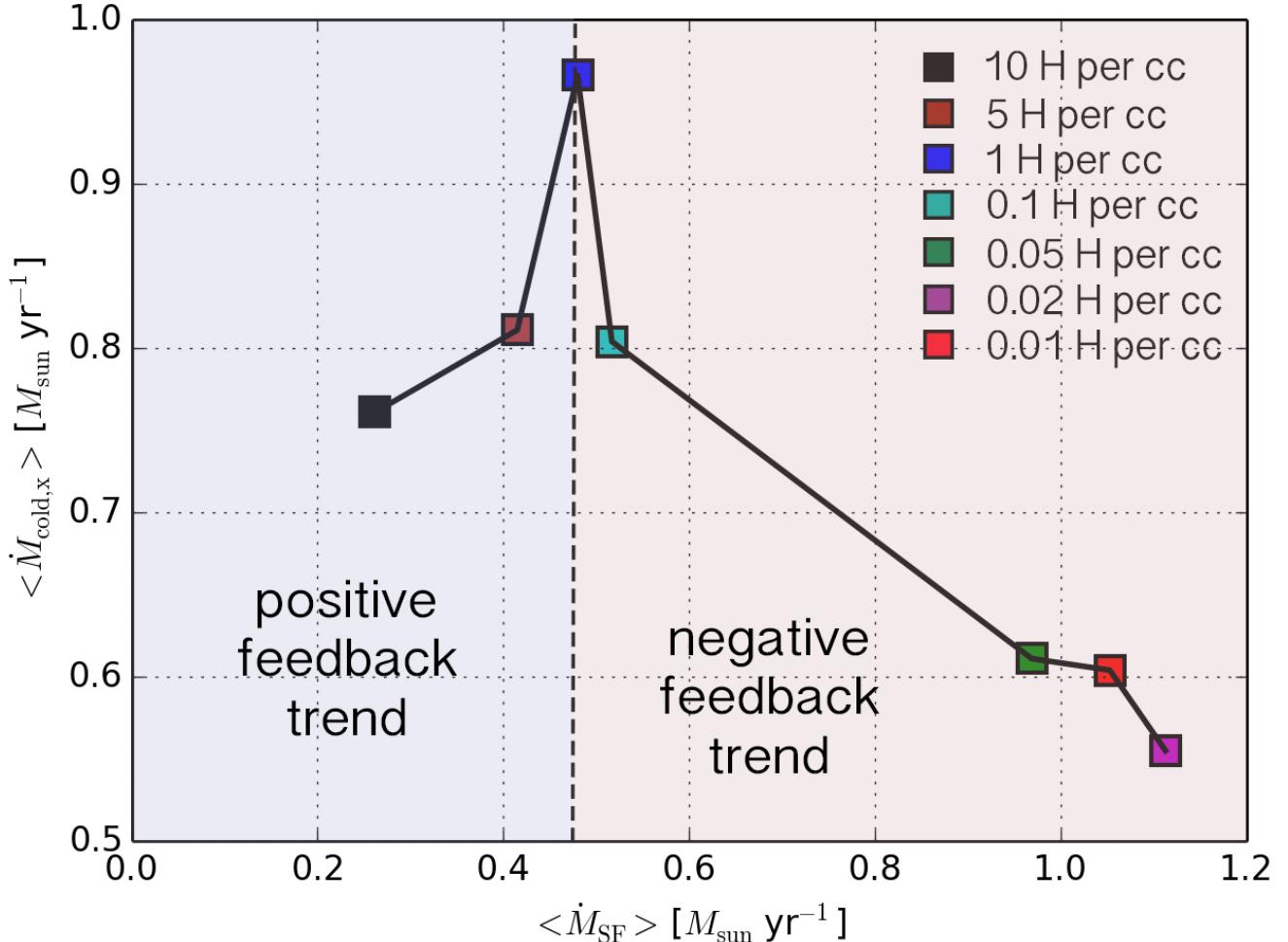
where  $n$  is the number density and  $\Lambda(T)$  is the cooling rate. The cooling time can be defined as

$$t_{\text{cool}} \equiv \frac{E}{|dE/dt|} = \frac{3nkT}{2n^2\Lambda(T)} = \frac{3kT}{2n\Lambda(T)} \quad (6)$$

where  $R$  is the gas constant. A parcel of initially  $10^6$  K gas interacting with an equal mass of  $10^4$  K gas will, therefore, assuming thermal equilibrium is reached (so that  $T_{\text{mixed}} \approx$

$5 \times 10^5$  K and the density is twice the initial density of the halo gas), decrease its cooling time by almost an order of magnitude. The cooling time for  $10^6$  K gas at 0.1 solar metallicity is  $\sim 0.5$  Gyr (Sutherland & Dopita, 1993), so for the mixture  $t_{\text{cool}} \sim 100$  Myr.

The mixing of the disc gas with low-metallicity halo gas is also an important factor in matching observations. Studies of G-type, K-type, and M-type dwarf stars (Schmidt, 1963Worthey et al., 1996Casuso & Beckman, 2004Woolf & West, 2012) in the MW show that they possess a cumulative metallicity distribution that peaks strongly at solar metallicity, with a low-Z tail falling off much faster than the predictions of 'closed-box' chemical evolution models, which find a Z-distribution that increases steadily with time and leads to a flatter slope. A solution to this discrepancy is near-continuous infall of metal-poor gas onto the Galactic disc, at a rate of  $\sim 1 M_{\odot} \text{ yr}^{-1}$ , in order to dilute the enrichment from stellar feedback in situ. The gas that is entrained in our model is of preferentially low metallicity, 0.1 solar in the IC, compared to the disc gas metallicity of 1 solar. As time goes on, some of this halo gas will increase in metallicity due to enrichment from SNe, but as can be seen in Figure 1 (bottom panel), low-Z gas from above and be-



**Figure 7.** Trend of extraction rate of cold gas mass with SFR, both averaged over 1 Gyr, for the H1-sf runs with SF thresholds as shown in the key. The low average SFRs show a trend of increasing average cold gas mass rates with increasing SFR, while the high SFRs show the opposite. We have therefore made a distinction between a ‘positive feedback trend’ and a ‘negative feedback trend’, with the peak being approx.  $0.5 \text{ M}_\odot \text{ yr}^{-1}$ .

low the immediate mixing region replaces the original low-Z gas that has mixed. It is clear, therefore, that this mechanism can bring metal-poor gas into the Galaxy. The mixing of the solar metallicity gas with the metal-poor gas has an added effect of increasing the mixture’s capacity to cool – simply mixing temperatures will bring the halo gas closer to the peak of the cooling curve, but mixing metallicities will move this peak higher in terms of cooling rate (Katz et al., 1996).

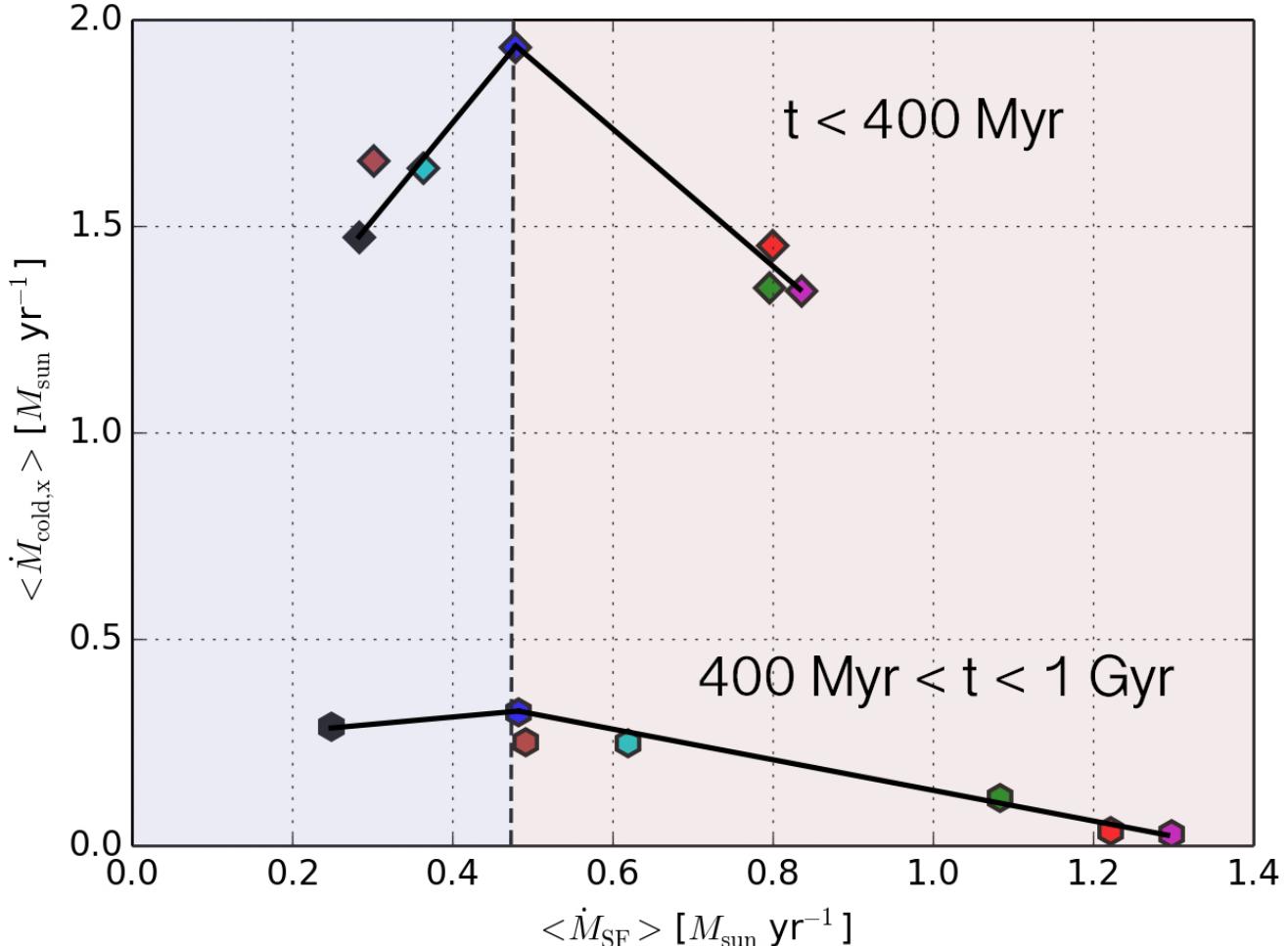
With the high resolution that we employ, and with the use of a physical thermal conductivity, it is likely that in certain places we start to resolve the Field length (Field, 1965; Armillotta et al., 2016) and see its effects on the evolution of the outflow. The Field length for cold gas structures embedded in a hot medium is given by,

$$\lambda_{\text{Field}} \equiv \sqrt{\left( \frac{\kappa_{\text{Sp}} T_{\text{hot}}}{n_{\text{cold}}^2 \Lambda(T_{\text{cold}})} \right)} \quad (7)$$

where  $\kappa_{\text{Sp}}$  is the Spitzer conductivity,  $n_{\text{cold}}$  and  $T_{\text{cold}}$  are the density and cooling rate of the cold gas respectively, and  $T_{\text{hot}}$  is the temperature of the hot ambient medium. For our

initial parameters, with the fiducial 0.1 Spitzer conductivity, the Field length is  $\sim 10$  pc. This quantity is important because it marks the boundary between a regime where thermal conduction dominates over radiative cooling, and vice versa.

For scales  $l \ll \lambda_{\text{Field}}$ , the temperature evolution of the cold gas structures moving through the hot medium is dominated by thermal conduction, whereas for  $l \gg \lambda_{\text{Field}}$  it is the radiative cooling that dominates. The outflowing cold gas that interacts with the hot gas can transition between the two regimes as it moves through the hot medium and breaks up, with the evolution of larger cold gas structures in the outflow dominated by radiative cooling but the late-time behaviour of the smaller remnants of the outflow being more heavily influenced at their boundaries by the thermal conduction from the hot gas. In our simulations we can take the smoothing length as the spatial resolution element, which varies adaptively over the course of the run but is typically  $\approx 10 - 20$  pc in the outflowing gas. In a few cases, gas within cold clumps at the edges of the outflow gets down to smoothing lengths of  $\approx 2 - 5$  pc. In these cases, therefore, we



**Figure 8.** Average rate of extracted cold gas mass trend with average SFR, for the H1-sf runs with SF thresholds as shown in the key in Figure 7. Here we take the average in two parts; for  $t < 400$  Myr (diamonds), and for  $400$  Myr  $< t < 1000$  Myr (pentagons). Note: lines connecting the symbols are purely instructive; to indicate an approximate trend (or not).

can be reasonably confident that thermal conduction with the hot gas influences the evolution of the smaller cold gas structures. This means that at any point in time there are two possible outcomes; either the outflow gas survives and increases in mass as it entrains more and more of the hot medium, or it evaporates in the hotter medium. The inclusion of a physical thermal conductivity in our simulations is therefore an important feature, as is the reasonable convergence of the results with varying  $\eta$  (see Figure 11).

#### 4.1 A constant cold gas mass through positive feedback

The SFR of the MW is almost flat over its lifetime (e.g., Rocha-Pinto et al., 2000), which implies a roughly constant cold gas reservoir from which stars can form. Within the framework of our current understanding of galaxy formation, maintaining such a reservoir is a challenge after  $z \sim 1$ , since the Galaxy no longer receives significant amounts of additional cold gas from its wider cosmological environment. The gas that is used up in star formation must somehow be replenished, however, in order to keep the cold gas mass in

the Galaxy approximately constant. As we have shown in Figure 2, with the positive feedback process extracting gas from the hot halo, we can maintain a roughly constant cold gas mass (blue line) for at least 3 Gyr (the limit of the fiducial H1 run). Of course, since this is a simplified model, the full picture is likely to be more complex, but it provides strong evidence for positive feedback helping to keep this cold gas reservoir constant.

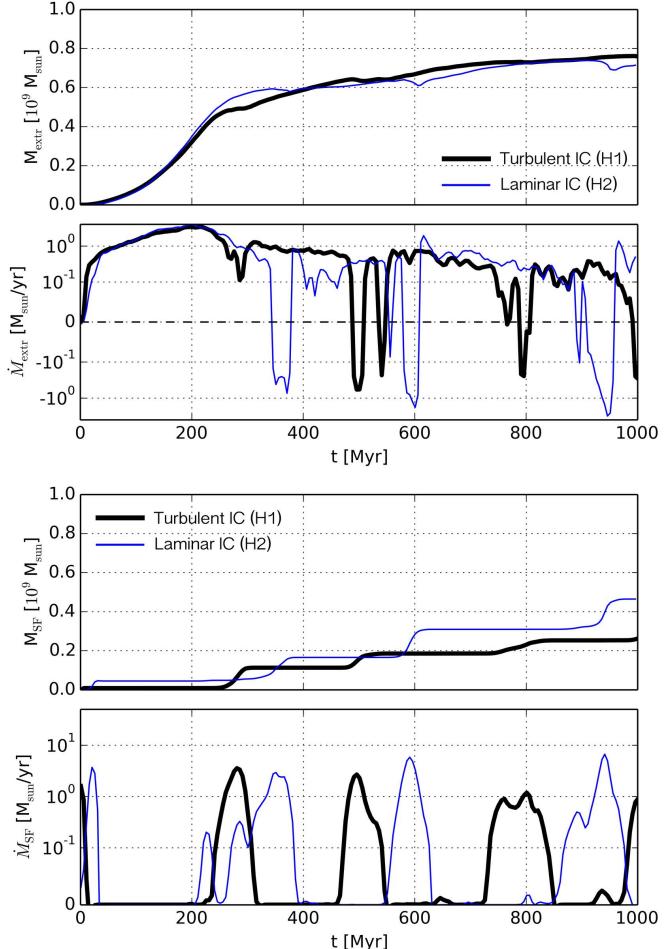
When we turn our attention to the *extracted* gas, rather than the total (Figure 4), we notice that the cold gas extracted from the halo and the hot gas extracted from the disc (in the latter case, by simple SNe heating) have very different histories. The extracted cold gas is on the whole fairly monotonic, with the  $\dot{M}$  staying mostly in the positive range, with the occasional dip below zero. The hot gas on the other hand, oscillates largely around zero in the  $\dot{M}$ , and ends up almost back where it started after 1 Gyr in the total mass (top plot). This makes sense when one considers the mechanism behind each mode – gas cooled from the halo will remain cold until it can form stars and/or be heated by SNe feedback, which happens mostly in the disc. Gas heated from the disc by SNe, however, will cool relatively quickly

**Table 2.**  $\dot{M}_{\text{cold}}$  is the average rate of ‘cold’ gas mass increase (defined as  $T < 2 \times 10^4 K$ ) and  $\dot{M}_{\text{hot}}$  is the average rate of hot gas mass increase (defined as  $T > 2 \times 10^5 K$ ) in the simulation over 1 Gyr.  $\dot{M}_{\text{SF}}$  is the average SF rate over 1 Gyr. The final two columns provide average mass conversion rates;  $\dot{M}_{\text{cold,x}}$  is the average rate of cold gas ( $T < 2 \times 10^4 K$ ) extracted from the initial hot ambient gas, and  $\dot{M}_{\text{hot,x}}$  is the average rate of hot gas ( $T > 2 \times 10^5 K$ ) formed from the disc gas. All values in this table are in units of  $M_{\odot} \text{ yr}^{-1}$ .

ID	$\dot{M}_{\text{cold}}$	$\dot{M}_{\text{hot}}$	$\dot{M}_{\text{SF}}$	$\dot{M}_{\text{cold,x}}$	$\dot{M}_{\text{hot,x}}$
H1	0.462	-0.873	0.262	<b>0.761</b>	0.051
H2	0.147	-0.791	0.465	<b>0.717</b>	0.130
H1-sf5	0.311	-0.886	0.415	<b>0.811</b>	0.105
H1-sf1	0.464	-1.095	0.480	<b>0.966</b>	0.059
H1-sf0.1	0.273	-0.951	0.517	<b>0.804</b>	0.052
H1-sf0.05	-0.432	-0.801	0.969	<b>0.611</b>	0.123
H1-sf0.02	-0.627	-0.743	1.113	<b>0.555</b>	0.129
H1-sf0.01	-0.517	-0.769	1.053	<b>0.604</b>	0.127

since the gas is dense. As such we see peaks in the hot gas  $\dot{M}$  at around the same times as the peaks in the SFR (Figure 6, black line) when the starbursts occur, but these hot gas peaks are typically short lived.

The runs with different SF thresholds (see Table 1) show a similar form of the extracted cold gas mass and the corresponding time derivative as in our fiducial H1 run. In Figure 5 we plot each on the same axes, and we find that the extracted cold gas mass is largely the same until  $\approx 250$  Myr, after which they start to diverge. The runs with an SF threshold above 0.1 atom per cc continue rising to 1 Gyr, while the runs with SF thresholds less than this flatten out (red, magenta, and green lines). In these latter two simulations, the SF threshold is so low that stars are able to form extremely easily, including in the cold gas that is sent out into the halo. Thus, the heating from SNe occurs not just each time the disc reforms but indeed across the whole duration of the simulation. This can be seen in Figure 6, where there is a clear difference between these two low SF threshold runs versus the others – the red, magenta, and green lines are largely continuous on the SFR plot (bottom panel), while the high SF threshold runs (black, blue lines) are bursty, only spiking when the disc has reformed enough to undergo another starburst. This burstiness gives us the time interval for our galactic fountain cycle, which is  $\approx 250 - 300$  Myr (for H1). Of course this differs for each run with different SFRs, since more powerful starbursts send the gas further out, causing it to take longer to return and re-form the disc.

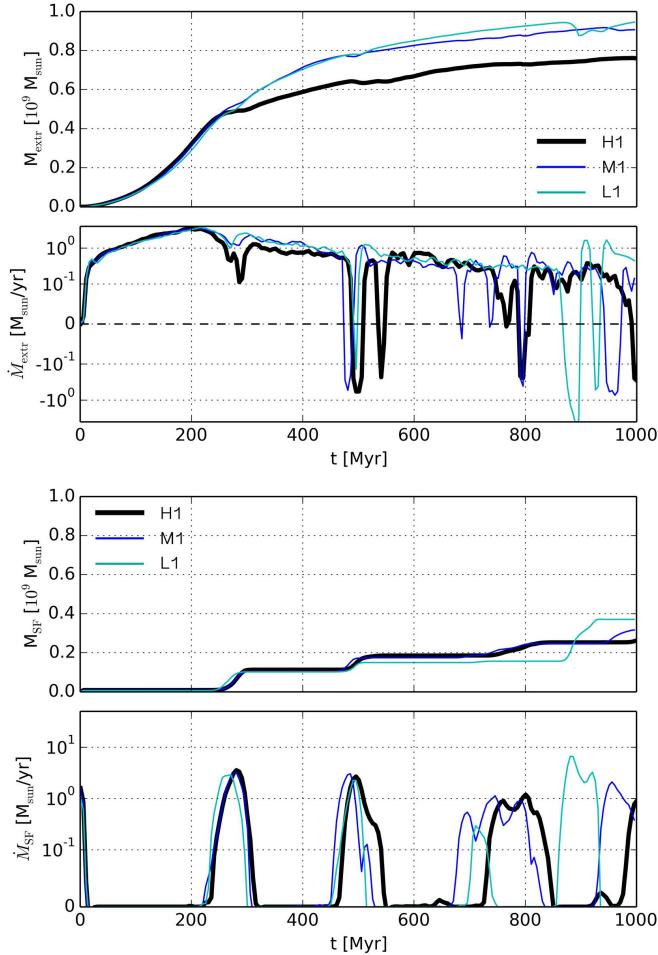


**Figure 9.** Comparison between extracted cold gas (top) and SF (bottom) for H1 and H2 - the fiducial run with the turbulent IC and the run with the non-turbulent IC.

#### 4.2 Positive feedback trend with SFR

In Table 2 we have recorded the averages of various mass rates over the first Gyr in units of solar masses per year for a number of the runs. The first thing to note about the numbers in this Table is due to the different definitions of ‘cold’ and ‘hot’ gas, the numbers for each run (along the row) should not be expected to add up to any constant value – instead, a more instructive comparison is down each column, where we can see how each average rate (cold gas, hot gas, SF, extracted cold gas, extracted hot gas) changes as we change the threshold density for SF. The SF column, in particular, shows us that as the SF threshold density decreases, the average SFR increases – this is not particularly surprising. The column highlighted in bold, however, which refers to the average mass rate of cold gas extracted from the halo gas, is more interesting in that the value increases in going from  $SF_{\text{thresh}} = 10$  to  $SF_{\text{thresh}} = 1$ , at which it peaks, and starts to decline to the lower SF threshold values.

Earlier in the paper we defined a ‘positive feedback trend’ to be increasing SFR  $\rightarrow$  increasing extracted cold gas, and the ‘negative feedback trend’ as vice versa. When we plot the average SFR vs. the average extracted cold gas rate (Figure 7), we see that we have two regimes – a positive



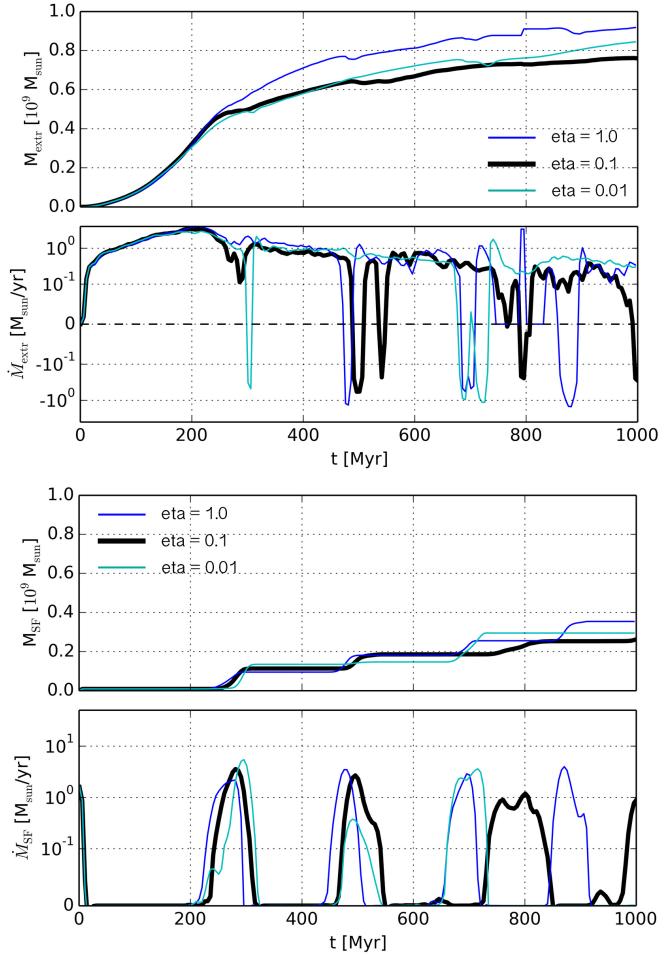
**Figure 10.** Comparison between three different resolutions – H1, M1 and L1 (see Table 1) for the extracted cold gas from the hot ambient medium (top) and the star formation history (bottom) over 1 Gyr.

feedback trend at low SFR values and a negative feedback trend at high SFR values, with the peak somewhere close to  $0.5 M_{\odot} \text{ yr}^{-1}$ .

We can construct a simple analytical model to describe the functional form of this plot. From the simulations we have performed, we can identify two reservoirs of gas; cold (the star-forming disc gas) and hot (the ambient medium). To encapsulate the positive and negative effects of stellar feedback on the cooling of hot halo gas discussed earlier, we introduce two (not necessarily constant) coefficients denoting positive ( $\epsilon_+$ ) and negative ( $\epsilon_-$ ) feedback efficiency<sup>1</sup>. Specifically,  $\epsilon_-$  includes the preventive effect of heating the hot medium by SNe which counteracts the cold gas extraction from the hot halo via positive feedback.

Our basic assumption then is that the extraction rate of cold gas is given by the net efficiency  $\epsilon = \epsilon_+ - \epsilon_-$  times

<sup>1</sup> Note that the latter is not the same as the ‘negative feedback’ defined in Figure 4, which was the gas from the cold disc that was heated by the SNe. Here we are only dealing with gas extracted from the hot medium.



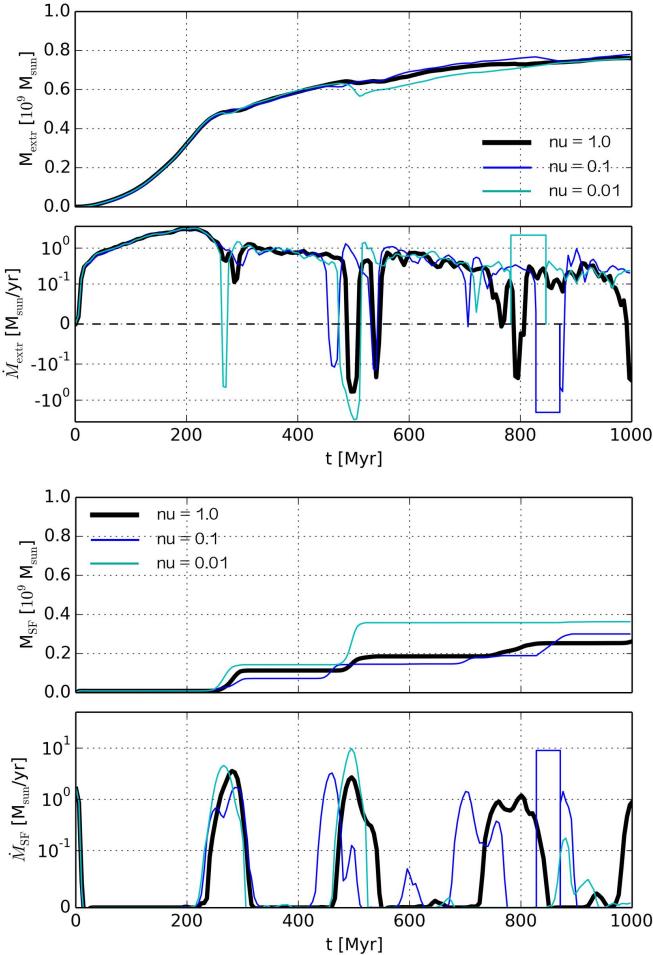
**Figure 11.** Comparison between three values of the thermal conductivity parameter,  $\eta = 1.0, 0.1, 0.01$  (fiducial) for the extracted cold gas from the hot ambient medium (top) and the star formation history (bottom) over 1 Gyr.

the SFR, i.e.,  $\dot{M}_{\text{cold},x} = (\epsilon_+ - \epsilon_-) \dot{M}_{\text{SF}}$ . We can write this as

$$\dot{M}_{\text{cold},x} = \epsilon_+ f \dot{M}_{\text{SF}}, \quad (8)$$

where  $f \equiv (1 - \epsilon_-/\epsilon_+)$  is, in general, a function of SFR. If we examine this function  $f$  we see that when  $\epsilon_- = \epsilon_+$ , the extraction of cold gas from the hot reservoir is completely opposed by the SNe reheating this gas, so that  $f = 0$ . We expect this scenario to occur at high SFRs. In contrast, when SFR are close to zero, we expect SNe heating to be sub-dominant compared to cold gas extraction, i.e.  $\epsilon_- \ll \epsilon_+$ . Hence, we expect  $f \sim 1$ . A simple assumption is thus  $f = e^{-\dot{M}_{\text{SF}}/\tau}$ . By taking the derivative and using the boundary between the two regimes as the stationary point (at  $\dot{M}_{\text{SF}} \approx 0.5$  – see Figure 7), we see that  $\tau \approx 0.5$ .

Furthermore, we make the assumption that  $\epsilon_+$  is approximately constant.  $\epsilon_+$  is the slope of the trend at low SFR (since  $f \approx 1$  at low SFR) and it measures the effectiveness of positive feedback for a given value of the SFR. Thus, a constant  $\epsilon_+$  implies that we will get the same mass of extracted cold gas for a given outflow mass. As a first approximation, this appears reasonable. Thus we have,



**Figure 12.** Comparison between three different values of the thermal conductivity parameter,  $\nu = 1.0$  (fiducial), 0.1, 0.01  $\nu_{\text{Brag}}$  for the extracted cold gas from the hot ambient medium (top) and the star formation history (bottom) over 1 Gyr. Note: horizontally-flat sections in the rate correspond to missing data.

$$\dot{M}_{\text{cold},x} = A e^{-\dot{M}_{\text{SF}}/\tau} \dot{M}_{\text{SF}} \quad (9)$$

where  $A \equiv \epsilon_+$  is the normalisation constant. We show this functional form in Figure 15, alongside the averaged data points from the simulations, for two (arbitrary) normalisations for purposes of illustration. We can see that the functional form of our simple analytic model fits the data points reasonably well, although the data points do not all lie on the same normalisation. Given that we adopted one of many potential choices for the functional form of  $f$ , our toy model may be a useful guide to understand the behavior of our simulations qualitatively, but not quantitatively.

This result, although tentative due to the nature of our simplified model, has interesting implications for the evolution of late-type galaxies. If we consider the different SFRs occurring not in different galaxies but in the *same* galaxy across its lifetime, then we can infer that while in the positive feedback trend regime (left part of the plot), any change to the SFR will be unstable – an increase in SFR will lead to a greater amount of cold gas extracted from the hot halo, which will in turn lead to more star formation and a further

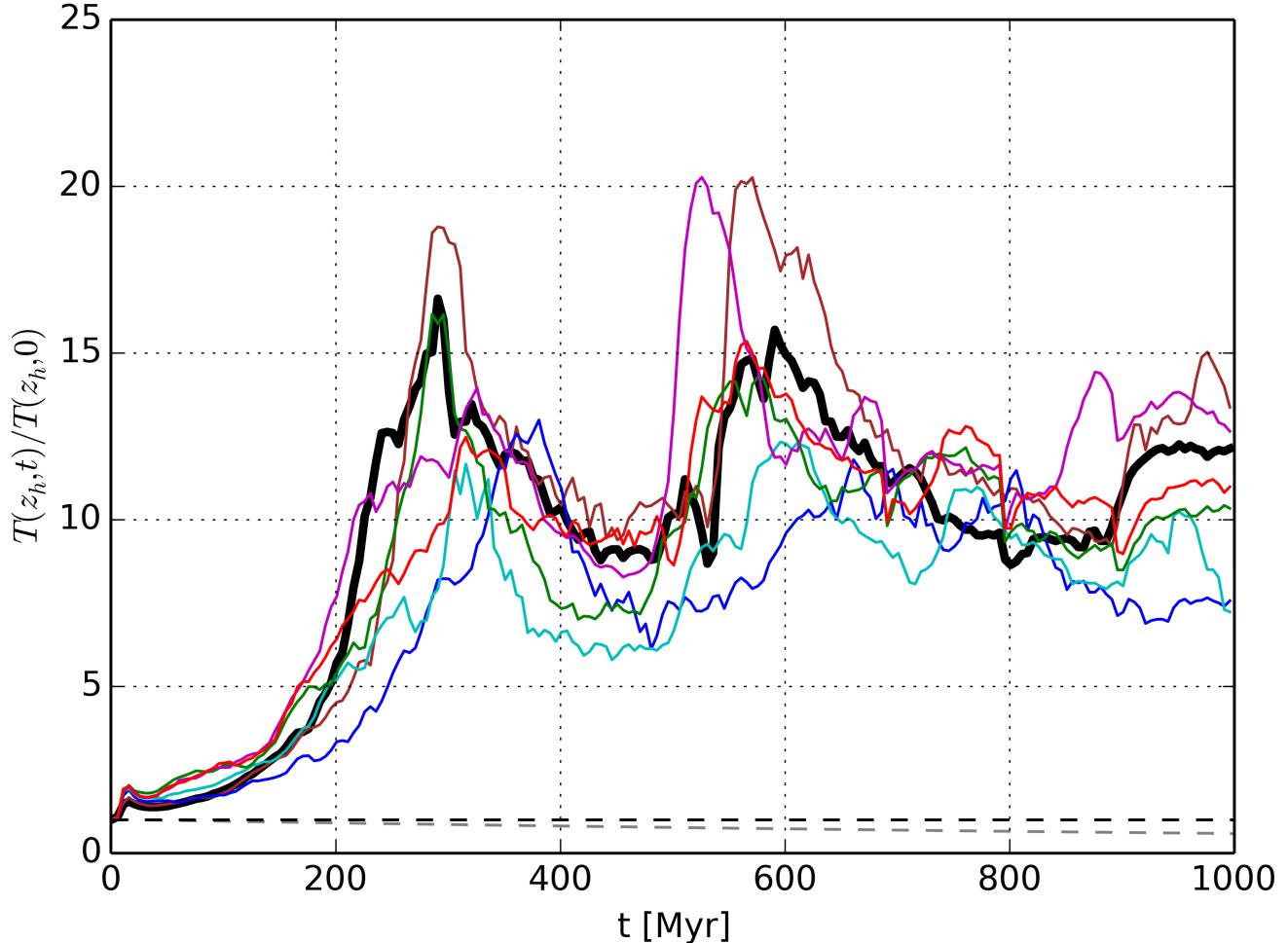
increase in SFR. Vice versa, if the SFR drops for a period of time, less cold gas will be extracted, causing the SFR to drop still further. As a result it is likely that galaxies would not stay where they are in this left-hand part of the plot, but rather reach either peak cold gas extraction (at  $\text{SFR} \approx 0.5 M_{\odot} \text{ yr}^{-1}$ ) or reach zero SFR ('red and dead'). These two values of the SFR are therefore 'attractors' in the sense that any perturbation to the SFR would cause it to attain one of these values. We note, of course, that the values in this plot are averaged SFRs over 1 Gyr, rather than instantaneous SFRs, and so we cannot comment on how quickly a galaxy would move left or right along this slope.

In the right-hand part of the plot (negative feedback trend), however, the behaviour of the galaxy subject to any perturbation of the SFR is very stable – an increase in SFR reduces the amount of new cold gas extracted from the halo, which subsequently causes the SFR to drop. Likewise, a decrease in the SFR causes it to rise again. Thus galaxies can in theory remain where they are on the negative feedback trend for a long time.

This simple stability discussion is depicted in cartoon schematic form in Figure 16. The peak of cold gas extraction rate that divides the two trends lies, in our suite of simulations, at  $\approx 1 M_{\odot} \text{ yr}^{-1}$ , which is in the approximate range for the required cold gas infall rate to solve the G-dwarf (and K- and M-dwarf) problem. The SFR for this peak is  $\approx 1 M_{\odot} \text{ yr}^{-1}$ , which is also in the ballpark range of the MW SFR over its lifetime. However, we stress that, due to the nature of our somewhat limited model, we do not wish to make any quantitative conclusions in this paper – it is entirely possible that even small modifications to our setup could result in the peak in the stability plot being shifted left or right in SFR and up or down in rate of cold gas extraction (not to mention that we currently lack error bars on these measurements, which might alter these numbers also). For now we prefer to speculate on the qualitative form of this plot rather than the precise normalisation.

#### 4.3 Effectiveness of positive feedback

In Armillotta et al. (2016), it was noted that the increase in cold gas mass as the cold cloud passed through the hot medium was strongly dependent on the temperature of the hot gas (Figure 6 in Armillotta et al., 2016). These authors found that an increase of only a factor of 8 from the initial temperature of  $10^6$  K reduced the amount of cold gas collected by the end of the 60 Myr period by a factor of  $\sim 30$ . To determine how a change in the ambient environment affects our positive feedback mechanism, we have plotted, against time, the ratio of the average temperature of the ambient medium at the approximate 'mixing height' at time  $t$  to the equivalent average temperature at that height at  $t = 0$ . This gives us the change in the average temperature of the hot gas (note: this is the 'initially' hot gas that is tagged at the beginning of the simulation) that the cold gas must try to entrain as time goes on. This is shown in Figure 13 for each SF threshold. It is clear that the heating from the SNe, being confined to a relatively narrow periodic box, quickly heats up the ambient gas above the disc where the outflow tries to mix with it. The temperature stays relatively close to  $\sim 10^6$  K for the first 200 Myr, but subsequently it is heated to



**Figure 13.** Plot showing the ratio of average temperature to initial temperature, for the hot ambient gas in a region between 0.5 and 1x the maximum height that the cold dense gas has reached by time  $t$ . Once again we compare simulations H1-sf with SF thresholds as listed in Figure 6. The fiducial H1 run is shown with a thick black line for clarity. We also include the average temperature of this region assuming ONLY the ambient halo gas (therefore any cooling is due strictly to radiative cooling losses), this is shown with a grey dashed line. The black dashed line marks the initial ratio of 1.

$\sim 10^7$  K and remains at around this temperature for the rest of the run.

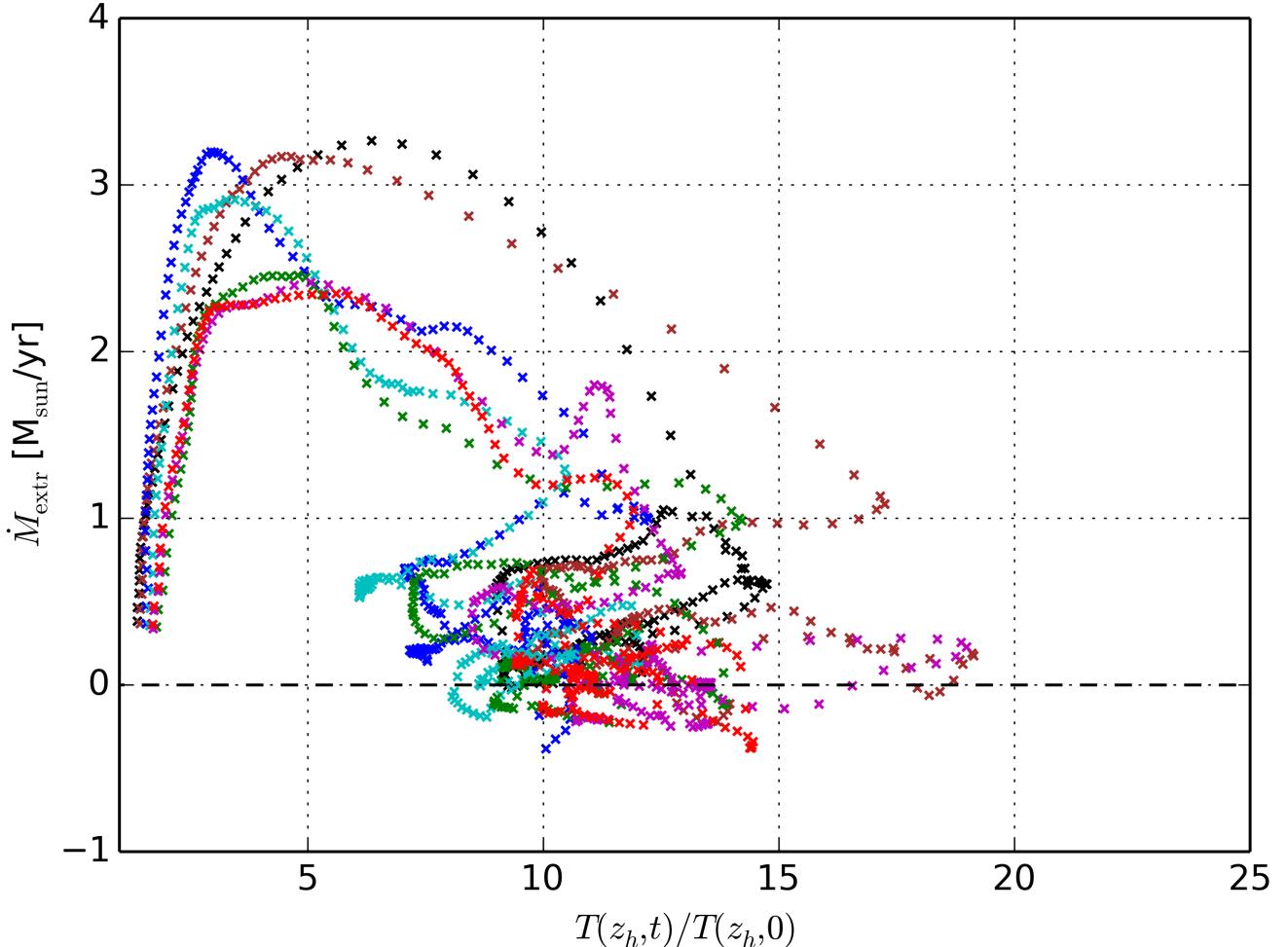
Based on this finding, and the finding of Armillotta et al. (2016), we find it logical to conclude that the reason for the reduction in rate of cold gas extraction after  $t \approx 300 - 400$  Myr is a hot medium that is just too hot for this process to be effective. We can plot a measure of this effectiveness by plotting our temperature ratio (y-axis of Figure 13) against our cold gas extraction rate (y-axis of Figure 5) as a running average with time to obtain Figure 14, which shows us that when the temperature is a few times  $10^6$  K, the outflow is effective at entraining cold gas, while at factors of 10–15 times the initial  $10^6$  K, this effectiveness drops substantially.

This is the reason that all of our runs show a substantial peak<sup>2</sup> in the cooling rate of the extracted gas (see Figure 4 for an example) around 200 Myr, corresponding to the

initial outflow after the first starburst occurs in the ICs. Subsequently to this peak, the rate of extraction of cold gas drops by a factor of  $\approx 5 - 10$  for the rest of the run. Given that this feature appears only within this initial period of time for every run, we explored the idea of treating this peak as an initial transient, that should be subtracted from our analysis, since perhaps the initial starburst is anomalous due to the setup of the ICs. However, upon further study we have concluded that this peak is not an effect of the initial, powerful starburst that should be subtracted, since as Figure 3 shows, the subsequent SFR peaks are roughly of equal magnitude to the initial one. *This shows that the initial starburst is no more powerful than the ones that come afterwards as a result of the disc reforming each time and continuing the cycle.*

Instead, it is clear that our initial setup provides the ‘optimal conditions’ for positive feedback to work. The gas hotter than a few  $\times 10^6$  K is no longer able to cool down to sub  $10^6$  K temperatures after mixing and thus retains both a high temperature and a long cooling time. Indeed,

<sup>2</sup> note that since the rate plots are plotted with a log y-axis, this strong peak may not be immediately apparent



**Figure 14.** Plot showing the temperature ratio from Figure 13 (the ratio of average temperature to initial temperature) vs. the extracted cold gas rate from Figure 5, to give an idea of how the efficiency of cold gas extraction varies with increasing ambient temperature. For this plot we take the running mean values on each axis over a time interval of 50 Myr. Note: crosses that are closer together in x-axis spacing indicate that the run spends more time at those temperatures.

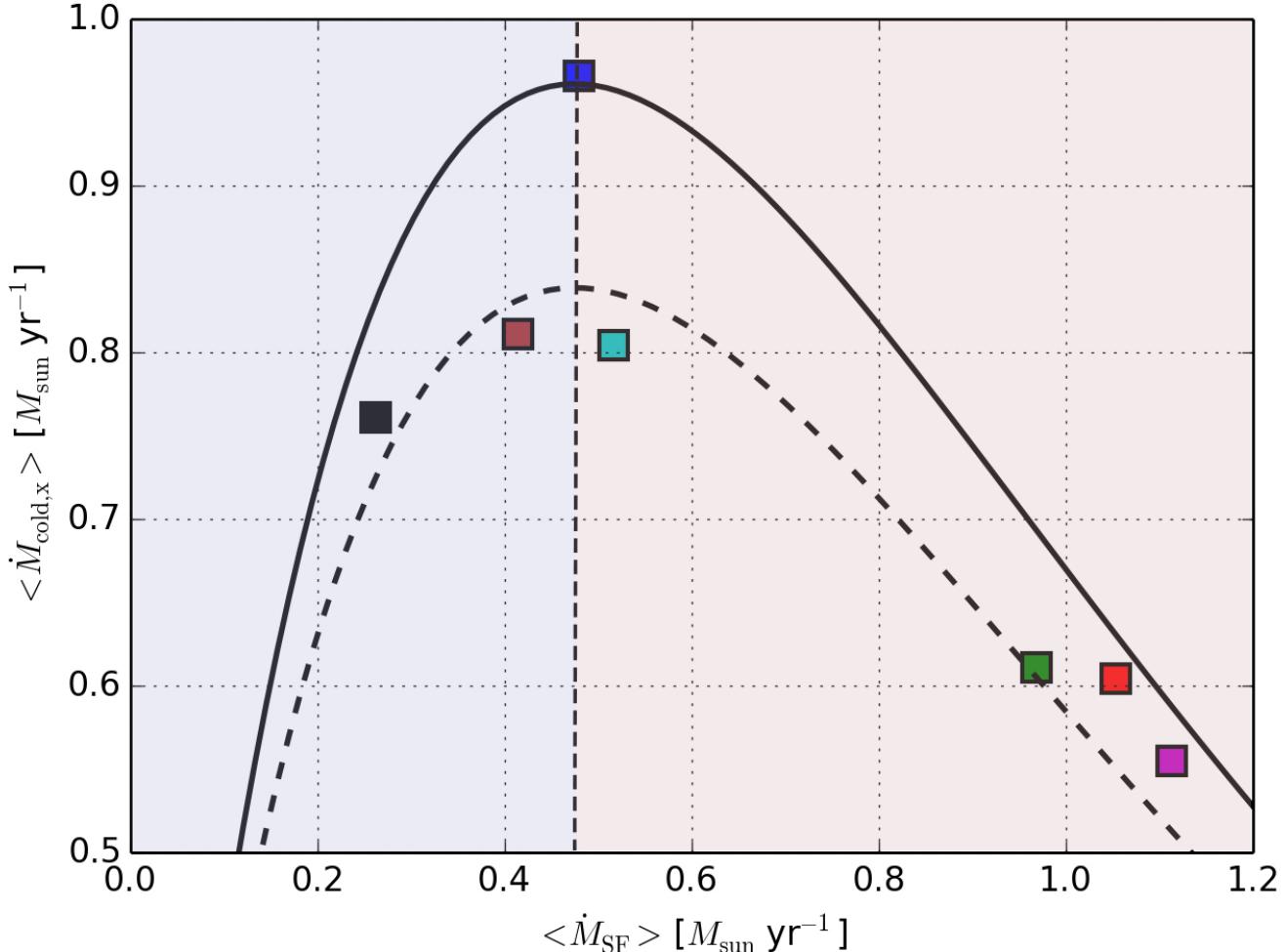
it is likely that the initial 300 – 400 Myr is a more realistic model of the positive feedback process than the subsequent period for two reasons: (1) our relatively simple (and low) SF threshold prescription allows stars to form and explode as SNe in the outflow, contributing to the heating of the ambient gas – this may not occur in real galaxies where the requirement for SF may be more stringent (2) the nature of the periodic boundaries in our setup causes the hot gas from the SNe to be ‘trapped’ in the domain rather than escaping. Any process in real galaxies that reduces the amount of hot gas from SNe in the mixing region will contribute to a far more efficient mode of positive feedback.

#### 4.4 An altered gas regulator model

We attempt to construct a simple schematic of how positive feedback might be included in a gas regulator model (e.g. Davé et al., 2012Lilly et al., 2013Feldmann, 2013Feld-

mann, 2015). In the usual picture, the cold<sup>3</sup> gas reservoir is increased by cosmological accretion and decreased by SF and outflows. In our picture we assume that cosmological accretion has ended, and instead invoke the hot halo as the source of gas that can grow the reservoir. This can be achieved either through radiative cooling ( $\dot{M}_\Lambda$ ) or through positive feedback ( $\dot{M}_+$ ). The sinks are the same as in other regulator models; gas being used up forming stars ( $\dot{M}_{SF}$ ) or ejected in outflows/negative feedback ( $\dot{M}_-$ ). This simple model is shown in Figure 17. The star formation now plays a dual role; through SNe feedback the SF can either give rise to negative feedback and help to deplete the reservoir, or it can give rise to positive feedback and help to grow the reservoir. Depending on the relative dominance of each mode, the reservoir will either decrease or increase respectively, which will lead to a similar decrease or increase in the SFR. This in turn will effect a change in the feedback

<sup>3</sup> although we note that most gas regulator models do not make a distinction on the temperature of the gas



**Figure 15.** Plot showing the functional form of our simple analytic model described by Equation 9 compared with the averaged data points from the simulations. We show two normalisations of the analytic expression, with  $A = 5.5$  (solid) and  $A = 4.8$  (dashed). Qualitatively, the functional form agrees quite well with the data points (see Section 4.2).

(both positive and negative), with an increase in SFR giving rise to more SNe. As discussed in Section 4.2, and shown in Figure 16, when averaged over 1 Gyr timescales, while the average SFR is below  $\approx 0.5 M_{\odot} \text{ yr}^{-1}$ , increasing the reservoir is likely to lead to more positive feedback dominance (which then further increases the reservoir and the SFR). Thus the reservoir and the SFR grow or shrink relatively quickly in this regime. Above the  $0.5 M_{\odot} \text{ yr}^{-1}$  value, however, negative feedback would likely start to dominate, but since it has a stabilising effect, the reservoir and SFR would reach approximately constant values.

This latter insight may have profound implications for galaxy evolution after  $z \sim 1$ . It is of course somewhat speculative, but if there indeed exists a slope  $\text{SFR} - \dot{M}_{\text{cold}}$  that is *increasing* up to a particular value, it is likely that this value would be reached quickly, should the SFR begin to increase for whatever reason. Conversely, if the opposite occurred, it is not unfeasible that the SFR would quickly drop all the way to near-zero, as the amount of cold gas returned to the galaxy via this positive feedback process became less and less. A slope of decreasing  $\text{SFR} - \dot{M}_{\text{cold}}$  would, on the other hand, mark a phase of galaxy evolution that was roughly

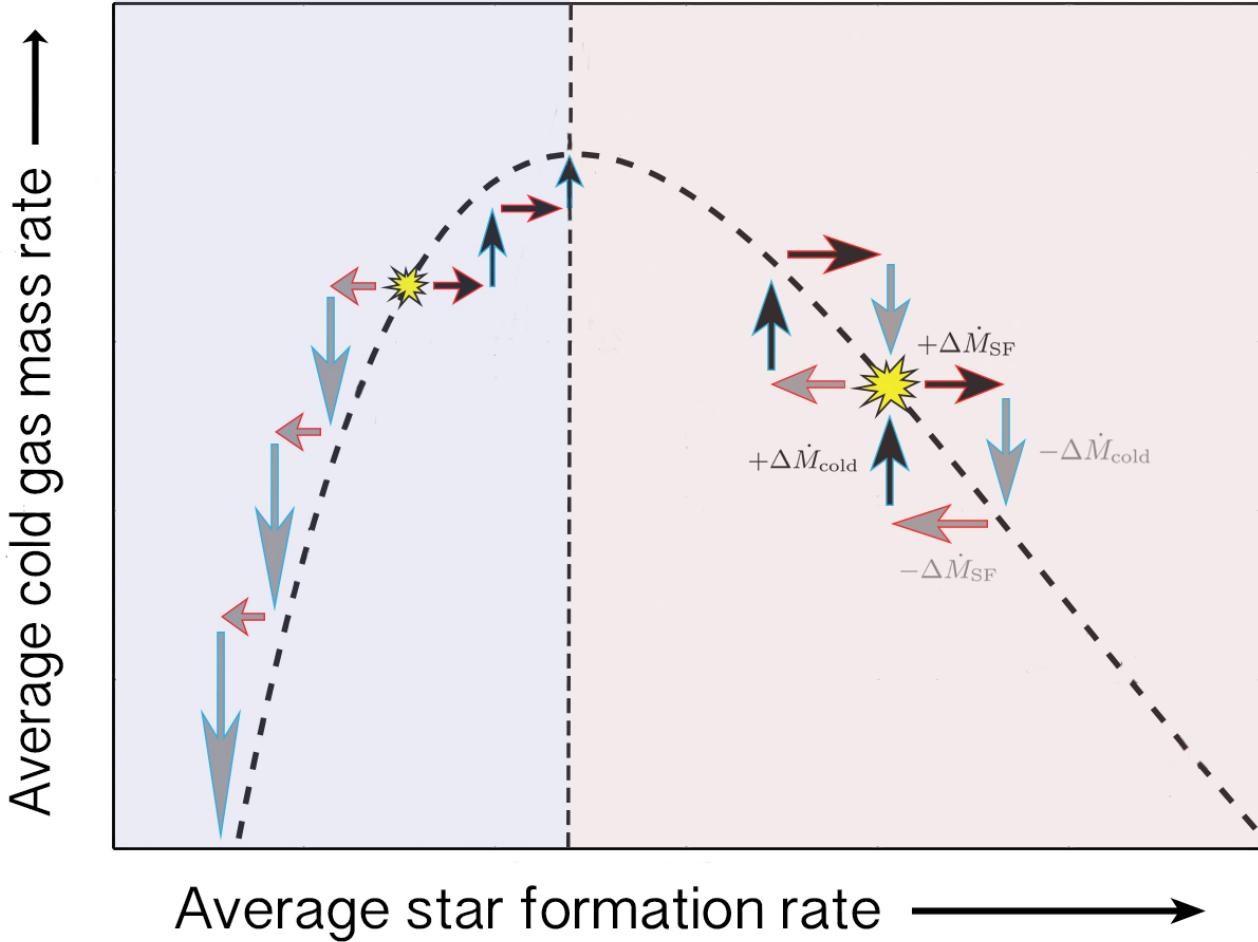
stable. This may have relevance to the standard ‘blue cloud – green valley – red sequence’ picture of SF quenching in galaxy formation.

The schematic shown in Figure 17 is of course merely one ‘sub-network’ of the larger galaxy formation network – it is the one relevant to our model, but the full picture would include contributions from, e.g., AGN feedback, cosmic rays, major and minor mergers, and of course cosmological accretion.

#### 4.5 Caveats

The simulations presented in this paper are simplified models, and many aspects of them are less than complete descriptions of the relevant physical processes. Here we list the relevant limitations of our model with respect to the real Universe.

*Star formation.* Our SF prescription is only based on a density threshold and molecular gas fraction (see Section 2). The formation of stars is therefore not dependant on whether the gas is Jeans unstable and as a result we may overestimate the number of stars formed. We stress however



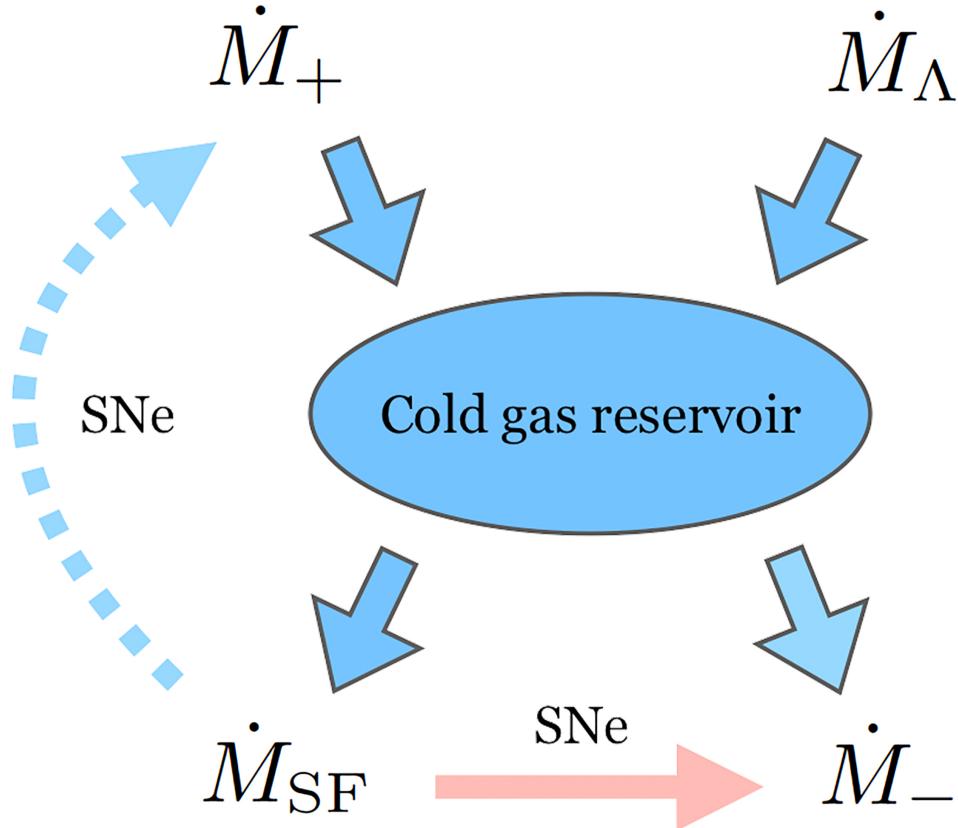
**Figure 16.** Cartoon demonstrating the stability of each side of the positive/negative feedback trend for a galaxy that lies on the trend line. Here we have used the analytical expression derived in Section 4.2, with a normalisation that is the average of the two normalisations shown in Figure 15. A galaxy that lies on the negative feedback trend (right-hand side) is very stable to perturbations in its average SFR. An increase (due perhaps to an external influence such as a merger or flyby) in average SFR would, assuming that the trendline is followed, lead to a decrease in the rate at which cold gas is extracted from the hot halo. This would in turn lead to a decrease in average SFR and a subsequent increase in the rate at which cold gas is extracted from the hot halo, bringing the galaxy back to where it started on the negative feedback trendline. Conversely, a perturbation that decreased the average SFR would lead to the same result but in the opposite direction. A galaxy that lies on the left-hand side of the plot, however (the positive feedback trend) would find any perturbation that either increased or decreased the SFR led to a continually increasing or decreasing extracted cold gas rate respectively. The position of the galaxy on this side of the trend would therefore be very unstable, tending to either increase in average SFR to the peak of the trend at  $\approx 0.5 M_{\odot}$  or decrease to zero.

that our model is not intended to be a fully realistic description of the Galaxy, but rather a means to explore the effect of different star formation rates on the existence and effectiveness of positive feedback.

*Boundary conditions.* The computational domain is restricted, using periodic boundaries in  $x$  and  $y$ , which allow us to sample the domain without losing any gas through these boundaries, but the effect of this is likely to increase the confining pressure of the gas within the domain, leading to a greater amount of heating as the outflow is not able to expand as it would in a full galaxy setup. In  $z$  we use an outflow boundary, which may introduce small numerical artifacts, although we have mitigated that by extending the vertical domain to 200 kpc in each direction.

*Initial conditions.* At the disc-ambient boundary of the ICs,

we implemented a minor numerical fix that prevented mixing due to spurious pressure forces within the first 200 Myr. These pressure forces are a natural consequence of having two different particle/cell distributions next to each other, even despite ensuring that the hydrodynamic pressure was constant across the boundary. The effect of these spurious forces was to cause hot gas from the ambient medium to sink onto the disc, and so the fix reduced the acceleration onto the disc in the vertical direction by a factor of 100, for the gas between 0.2 and 0.4 kpc above the disc plane. After 200 Myr, during the first outflow phase of the galactic fountain cycle, the disc and ambient gas distributions had mixed sufficiently that the numerical correction was no longer needed. The effect of our correction is to actually reduce mixing at the disc-ambient boundary for the first 200 Myr, so we do



**Figure 17.** Cartoon schematic for a low-redshift (i.e., assuming no cosmological accretion) gas regulator model that includes positive feedback. The sources are radiative cooling ( $\dot{M}_\Lambda$ ) and the positive feedback cold gas extraction mechanism outlined in this paper ( $\dot{M}_+$ ) and the sinks are star formation ( $\dot{M}_{SF}$ ) and negative feedback i.e., heating from SNe ( $\dot{M}_-$ ). Both the positive and negative feedback are caused by SF via the SNe, although in the case of the positive feedback this is more indirect. It should be noted that in this schematic we make no comment on the relative quantitative contribution of each source or sink.

not expect it to have any bearing on our conclusions.

*Velocity profile.* We model the hot halo without a shear velocity that would account for the difference in rotational velocity between the disc and the halo. It is likely that including the correct rotational profile of the gas would serve to re-accrete the outflow gas faster onto the disc and change the periodicity of the Galactic fountain cycle in our model. We base this on work by Fraternali & Binney (2008), who simulated galactic fountain clouds moving in the potentials of NGC 891 and NGC2403, where the modelled interaction of these clouds with the hot ambient gas was necessary in order to reproduce the vertical rotational lag. These authors found that the process of entraining gas from a hot halo with lower rotational velocity than the galactic disc caused the clouds to lose angular velocity and promoted their accretion onto the disc.

## 5 CONCLUSIONS

We have presented the results of a numerical model that seeks to understand the role of SNe outflows in increasing or decreasing the cold gas content of the MW (and by ex-

tension, other late-type spiral galaxies). We have employed a simplified setup of a  $2 \text{ kpc} \times 2 \text{ kpc}$  portion of the MW disc (present day) that is allowed to form stars which subsequently explode as SNe. From the series of numerical experiments we have conducted using this model we have found the following:

- (i) Feedback from SNe can be both positive AND negative in the context of promoting or hindering SF in the Galaxy.
- (ii) The action of SNe feedback giving rise to an outflow of cold ( $T < 2 \times 10^4 \text{ K}$ ) gas causes additional cold gas to be collected from the originally hot ambient medium (presumed to be the hot halo gas surrounding the galaxy).
- (iii) SNe also heat up cold gas, reducing the supply, but their role in creating an outflow that cools and entrain the hot gas dominates over their role in heating cold gas.
- (iv) This process of positive feedback is strongly dependent on the temperature of the hot medium from which the gas is cooled and entrained, with  $T \sim 10^6 \text{ K}$  allowing for an efficient positive feedback mechanism but  $T \sim 10^7 \text{ K}$  reducing the effectiveness considerably.
- (v) The ability of SNe outflows to increase the amount of cold gas in a galaxy may improve with increasing SFR, but

only up to a point – in our case,  $0.5 M_{\odot} \text{ yr}^{-1}$  – after which negative feedback dominates.

(vi) The typical values for the rate of extraction of cold gas in our simulations ( $0.5 - 1.0 M_{\odot} \text{ yr}^{-1}$ ) is in approximate agreement with the required rate for the MW to maintain its near-constant SFR since  $z \sim 1$ .

(vii) The presence of a peak in the average SFR vs. average extracted cold gas plot, for which we have constructed a simple analytic model, has interesting implications (see Section 4) for galaxy evolution after  $z \sim 1$  when positive feedback may start to dominate the gas budget – is this peak value an ‘attractor’ that galaxies tend to reach as their SFR increases through the action of positive feedback? Is it just as likely that the SFR would decrease and lead to a rapid reduction of both positive feedback and SFR until the galaxy was ‘red and dead’?

Our results are a promising addition to the new but growing field of positive feedback in galaxy formation. We have drawn on the work of Armillotta et al. (2016), and related papers, and demonstrated that the effect they observe in simulations of individual cold clouds moving through hot gas can also take place in the larger, galactic outflow context. Subsequently, we have demonstrated that such a positive feedback mechanism likely varies with average SFR in a way that places a galaxy undergoing this process in one of two evolution regimes. We plan to conduct future work in this area by (i) running high-resolution models of whole (isolated) galaxies and (ii) attempting to detect this process of positive feedback occurring in state-of-the-art cosmological zoom-in simulations such as those conducted by the FIRE collaboration (Hopkins et al., 2014).

## 6 ACKNOWLEDGMENTS

The authors thank the referee for multiple suggestions that helped to improve the quality of the paper. AH thanks Pedro Capelo and Simon Lilly for useful insights, and Steffi Walch for the SILCC papers that inspired parts of this work. RF acknowledges financial support from the Swiss National Science Foundation (grant no 157591). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project IDs s698, s797, s926 and uzh18.

## REFERENCES

- Anderson M. E., Churazov E., Bregman J. N., 2016, MNRAS, 455, 227
- Armillotta L., Fraternali F., Marinacci F., 2016, MNRAS, 462, 4157
- Babyl I. V., McNamara B. R., Nulsen P. E. J., Russell H. R., Vantyghem A. N., Hogan M. T., Pulido F. A., 2018, ApJ, 862, 39
- Bogdán Á., et al., 2013, ApJ, 772, 97
- Bregman J. N., 1980, ApJ, 236, 577
- Casuso E., Beckman J. E., 2004, A&A, 419, 181
- Cavagnolo K. W., Donahue M., Voit G. M., Sun M., 2008, ApJ, 683, L107
- Cicone C., et al., 2018, ApJ, 863, 143
- Colbrook M. J., Ma X., Hopkins P. F., Squire J., 2017, MNRAS, 467, 2421
- Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98
- Dekel A., Mandelker N., 2014, MNRAS, 444, 2071
- Dubinski J., Narayan R., Phillips T. G., 1995, ApJ, 448, 226
- Feldmann R., 2013, MNRAS, 433, 1910
- Feldmann R., 2015, MNRAS, 449, 3274
- Fernández X., Joung M. R., Putman M. E., 2012, ApJ, 749, 181
- Field G. B., 1965, ApJ, 142, 531
- Fraternali F., Binney J. J., 2008, MNRAS, 386, 935
- Fraternali F., Tomassetti M., 2012, preprint, ([arXiv:1207.0093](https://arxiv.org/abs/1207.0093))
- Gaensler B. M., Madsen G. J., Chatterjee S., Mao S. A., 2008, doi:10.1071/AS08004, 25, 184
- Grcevich J., Putman M. E., 2009, ApJ, 696, 385
- Gronke M., Oh S. P., 2018, MNRAS, 480, L111
- Haardt F., Madau P., 2012, ApJ, 746, 125
- Hippelein H., et al., 2003, A&A, 402, 65
- Hobbs A., Read J., Power C., Cole D., 2013, MNRAS, 434, 1849
- Hobbs A., Read J., Nicola A., 2015, MNRAS, 452, 3593
- Hobbs A., Read J. I., Agertz O., Iannuzzi F., Power C., 2016, MNRAS, 458, 468
- Hogan M. T., et al., 2017, ApJ, 851, 66
- Hopkins P. F., 2015, MNRAS, 450, 53
- Hopkins P. F., 2017a, preprint, ([arXiv:1712.01294](https://arxiv.org/abs/1712.01294))
- Hopkins P. F., 2017b, arXiv e-prints, p. arXiv:1712.01294
- Hopkins P. F., 2017c, MNRAS, 466, 3387
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
- Hopkins P. F., et al., 2018, MNRAS, 477, 1578
- Joung M. R., Bryan G. L., Putman M. E., 2012, ApJ, 745, 148
- Kacprzak G. G., Churchill C. W., Steidel C. C., Murphy M. T., 2008, AJ, 135, 922
- Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19
- Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
- Kim J.-h., et al., 2014, ApJS, 210, 14
- Krumholz M. R., Gnedin N. Y., 2011, ApJ, 729, 36
- Leitner S. N., Kravtsov A. V., 2011, ApJ, 734, 48
- Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, ApJL, 460, L1
- Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119
- Madau P., Pozzetti L., Dickinson M., 1998, ApJ, 498, 106
- Magnani L., Smith A. J., 2010, ApJ, 722, 1685
- Marasco A., Fraternali F., 2011, A&A, 525, A134
- Marasco A., Fraternali F., Binney J. J., 2012, MNRAS, 419, 1107
- Marinacci F., Fraternali F., Nipoti C., Binney J., Ciotti L., Londrillo P., 2011, MNRAS, 415, 1534
- Mastropietro C., Moore B., Mayer L., Wadsley J., Stadel J., 2005, MNRAS, 363, 509
- Melioli C., Brighenti F., D’Ercole A., de Gouveia Dal Pino E. M., 2008, MNRAS, 388, 573
- Miller M. J., Bregman J. N., 2015, ApJ, 800, 14
- Nelson D., Vogelsberger M., Genel S., Sijacki D., Kereš D., Springel V., Hernquist L., 2013, MNRAS, 429, 3353
- Nicastro F., Mathur S., Elvis M., 2008, Science, 319, 55
- Puchwein E., Springel V., 2013, MNRAS, 428, 2966
- Putman M. E., Peek J. E. G., Joung M. R., 2012, ARA&A, 50, 491
- Rafferty D. A., McNamara B. R., Nulsen P. E. J., 2008, ApJ, 687, 899
- Rocha-Pinto H. J., Scalo J., Maciel W. J., Flynn C., 2000, ApJL, 531, L115
- Röhser T., Kerp J., Lenz D., Winkel B., 2016, A&A, 596, A94
- Sancisi R., Fraternali F., Oosterloo T., van der Hulst T., 2008, AAPR, 15, 189
- Schmidt M., 1963, ApJ, 137, 758
- Shapiro P. R., Field G. B., 1976, ApJ, 205, 762

- Spitzer Jr. L., 1942, ApJ, 95, 329  
Spitzer L., 1962, Physics of Fully Ionized Gases  
Stinson G. S., Dalcanton J. J., Quinn T., Kaufmann T., Wadsley  
J., 2007, ApJ, 667, 170  
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253  
Walch S., et al., 2015, MNRAS, 454, 238  
Wilson C. D., 2013, in Wong T., Ott J., eds, IAU Symposium Vol.  
292, Molecular Gas, Dust, and Star Formation in Galaxies. pp  
119–126, doi:10.1017/S174392131300077X  
Woods R. M., Wadsley J., Couchman H. M. P., Stinson G., Shen  
S., 2014, MNRAS, 442, 732  
Woolf V. M., West A. A., 2012, MNRAS, 422, 1489  
Worthey G., Dorman B., Jones L. A., 1996, AJ, 112, 948