



HAL
open science

NIHAO XIX: how supernova feedback shapes the galaxy baryon cycle

Édouard Tollet, Andrea Cattaneo, Andrea V. Macciò, Aaron A. Dutton, Xi Kang

► **To cite this version:**

Édouard Tollet, Andrea Cattaneo, Andrea V. Macciò, Aaron A. Dutton, Xi Kang. NIHAO XIX: how supernova feedback shapes the galaxy baryon cycle. *Monthly Notices of the Royal Astronomical Society*, 2019, 485, pp.2511-2531. <10.1093/mnras/stz545>. <insu-03718076>

HAL Id: insu-03718076

<https://insu.hal.science/insu-03718076v1>

Submitted on 9 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

NIHAO XIX: how supernova feedback shapes the galaxy baryon cycle

Édouard Tollet,^{1,2★} Andrea Cattaneo,¹ Andrea V. Macciò^{1b},^{3,4} Aaron A. Dutton^{1b}³ and Xi Kang⁵

¹Observatoire de Paris, GEPI and LERMA, PSL University, 61 Avenue de l'Observatoire, F-75014 Paris, France

²Université Paris Diderot, Sorbonne Paris Cité, F-75013 Paris, France

³New York University of Abu Dhabi, P.O. Box 129188, Saadiyat Island, Abu Dhabi, UAE

⁴Max Planck Institute für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

⁵Purple Mountain Observatory, the Partner Group of MPI für Astronomie, 2 West Beijing Road, Nanjing 210008, China

Accepted 2019 February 20. Received 2019 February 11; in original form 2018 November 14

ABSTRACT

We have used the NIHAO (Numerical Investigation of a Hundred Astrophysical Objects) simulations to explore how supernovae (SNe) affect star formation in galaxies. We find that SN feedback operates on all scales from the interstellar medium (ISM) to several virial radii. SNe regulate star formation by preventing condensation of HI into H₂ and by moving cold neutral gas to the hot HII phase. The first effect explains why the cold neutral gas in dwarf galaxies forms stars inefficiently. The second maintains the hot ISM of massive galaxies (HII vents out at lower masses). At $v_{\text{vir}} \gtrsim 67 \text{ km s}^{-1}$, the outflow rate follows the relation: $\dot{M}_{\text{out}} = 23 (v_{\text{vir}}/67 \text{ km s}^{-1})^{-4.6} \text{ SFR}$. 20 per cent–70 per cent of the gas expelled from galaxies escapes from the halo (ejective feedback) but outflows are dominated by cold swept-up gas, most of which falls back on to the galaxy on a $\sim 1 \text{ Gyr}$ time-scale. This ‘fountain feedback’ reduces the masses of galaxies by a factor of 2–4, since gas spends half to three quarter of its time in the fountain. Less than 10 per cent of the ejected gas mixes with the hot circumgalactic medium and this gas is usually not reaccreted. On scales as large as $6r_{\text{vir}}$, galactic winds divert the incoming gas from cosmic filaments and prevent it from accreting on to galaxies (pre-emptive feedback). This process is the main reason for the low baryon content of ultradwarves.

Key words: supernovae: general – ISM: jets and outflows – galaxies: evolution – galaxies: formation.

1 INTRODUCTION

Matching the abundances of galaxies and dark matter (DM) haloes (Leauthaud et al. 2012; Papastergis et al. 2012; Behroozi, Wechsler & Conroy 2013; Moster, Naab & White 2013; Tollet et al. 2017), the Tully–Fisher relation (Geha et al. 2006; Pizagno et al. 2007; Springob et al. 2007) and the rotation curves of spiral galaxies (Persic, Salucci & Stel 1996) show that a halo of mass $M_{\text{vir}} \sim 10^{11} M_{\odot}$ contains a galaxy with a typical stellar mass of $M_{\star} \sim 3 \times 10^8 M_{\odot}$. Many dwarf galaxies are rich in cold neutral gas due to inefficient star formation (e.g. Boselli et al. 2014; also see the simulations by Munshi et al. 2013 and Hopkins et al. 2014). Its presence can push the baryonic galaxy mass up to $M_{\text{gal}} \sim 10^9 M_{\odot}$ (Papastergis et al. 2012). However, even with gas, the galaxy mass is still an order of magnitude lower than the one expected from the universal baryon fraction $f_b = 0.16$.

The most common explanation is that the rest of the baryons have been blown out by supernova (SN) feedback (Dekel & Silk 1986).

However, a purely ejective scenario requires mass-loading factors of $\eta \sim 50$. For $1 M_{\odot}$ that forms stars, $50 M_{\odot}$ of gas need to be expelled, and the wind speed must be larger than the escape speed from the halo to prevent gas from coming back. These conditions are not easily met. Observed stellar outflows have typical mass-loading factors of order unity (Bouché et al. 2012; Cicone et al. 2014; Schroetter et al. 2015, 2016) if not smaller (Kacprzak et al. 2014; Pereira-Santaella et al. 2016; Falgarone et al. 2017).

Hydrodynamic simulations, too, have for a long time struggled to produce systems with mass-loading factors $\eta \gg 1$. Only in the present decade have they begun to form realistic disc galaxies and to reproduce the observed $M_{\star}-M_{\text{vir}}$ relation (Guedes et al. 2011; Brook et al. 2012; Hopkins et al. 2014; Shen et al. 2014; Vogelsberger et al. 2014; Schaye et al. 2015; Crain et al. 2015; Wang et al. 2015).

Simulations by Brook et al. (2014), Nelson et al. (2015), and Mitchell et al. (2018) have shown that the combined action of heating by a photoionizing ultraviolet (UV) background and outflows can prevent gas from falling into the shallow gravitational potential wells of haloes with $M_{\text{vir}} \ll 10^{11} M_{\odot}$. However, galaxies with $M_{\text{vir}} \gtrsim 10^{11} M_{\odot}$ do accrete a baryon mass of order $f_b M_{\text{vir}}$ and expel from the halo less than half that value. Hence, massive spirals

* E-mail: edouard.tollet@obspm.fr

must expel a lot of gas that remains within the virial radius. This gas is potentially available for reaccretion, and Brook et al. (2014) find that almost 40 per cent of the stellar mass at $z = 0$ is composed of reaccreted baryons.

Brook et al.'s (2014) 40 per cent figure is based on defining outflows as gas that flows out of a spherical surface of radius $r \sim 0.1r_{\text{vir}}$. Nelson et al. (2015), Anglés-Alcázar et al. (2017), and Mitchell et al. (2018) adopted similar criteria.¹ However, gas may be heated by SNe, vent out of galactic discs, cool, and fallback on to galaxies without crossing this surface, so that the real contribution of ejected and reaccreted baryons may be even larger.

Christensen et al. (2016) added a density and temperature criteria to decide which gas particles belong to galaxies, and they found that between 20 per cent and 70 per cent of the ejected gas was later reaccreted. By tracking gas particles, they could also measure the average reaccretion time-scale, which they found to be of the order of 1 Gyr. However, they did not follow the evolution of the ejected gas in density and temperature.

Our work is based on a Numerical Investigation of a Hundred Astrophysical Objects (NIHAO, Wang et al. 2015). Besides the size our sample, its novelty is that we track the motions of gas particles in both physical and thermodynamic phase space (that is, on a density–temperature diagram). Observed outflows have a multiphase structure with different speeds for gas that emits in CO, H α , and X-rays (e.g. McKeith et al. 1995; Lehnert, Heckman & Weaver 1999; Strickland & Heckman 2009; Beirão et al. 2015; Leroy et al. 2015 for the winds in M82). Our approach allows us to separate the different thermodynamic phases of simulated outflows and to follow their individual reaccretion time-scales. We shall show that this separation is important for a meaningful comparison with outflow rates measured from observations.

The primary goal of our study is to disentangle the relative contributions of the different mechanisms and pathways through which stellar feedback reduces the stellar masses of galaxies, since stellar feedback is not only a multiphase but also a multiscale phenomenon. Our study has also the secondary goal to improve the description of feedback in semi-analytic models (SAMs) of galaxy formation.

The article begins by presenting the NIHAO simulations (Section 2) and by describing the procedure used to assign particles to the different thermodynamic phases present within haloes (Section 3). Then, it presents our results for the galaxy baryon cycle, which we separate into accretion (Section 4), star formation (Section 5), ejection (Sections 6 and 7), and recycling (Section 8). Section 9 summarizes the conclusions of the article.

2 NIHAO SIMULATIONS

The simulations used for this work are from Wang et al.'s (2015) NIHAO. They were run with cosmological parameters from the Planck Collaboration et al. (2014): Hubble constant $H_0 = 67.1 \text{ km s}^{-1} \text{ Mpc}^{-1}$, matter density $\Omega_m = 0.3175$, dark-energy density $\Omega_\Lambda = 0.6824$, baryon density $\Omega_b = 0.049$, power-

¹ Anglés-Alcázar et al. (2017) use two stellar half-mass radii, which contain 80 per cent of the total stellar mass on average. For most galaxies, this criterion gives a $r < 0.1r_{\text{vir}}$. This may explain why Anglés-Alcázar et al. (2017) find higher recycled fractions (60 per cent–90 per cent) than both Brook et al. (2014) and Christensen et al. (2016).

Table 1. Best-fitting parameters for equation (7) at $z = 0$.

η	η_0	α
η_{coldISM}	7	-3.2 ± 0.3
η_{gal}	4	-4.6 ± 0.5
$\eta_{\text{gal}}^{\text{cold}}$	2	-5.6 ± 0.5
η_{halo}	2	-3.5 ± 0.5

spectrum normalization $\sigma_8 = 0.8344$, and power-spectrum slope $n = 0.9624$.

From three cubic N -body volumes of side lengths 15, 20, and $60 h^{-1} \text{ Mpc}$ (Dutton & Macciò 2014), Wang et al. (2015) selected ~ 90 haloes with masses ranging from $10^{9.5}$ to $10^{12.3} M_\odot$ discarding all haloes with a companion of comparable mass within $3r_{\text{vir}}$ at any time. The chosen haloes were resimulated at higher resolution with baryons.

At the highest level of refinement, the force softening length is $\simeq 50 - 80 \text{ pc}$ for the gas particles and $\simeq 100 - 200 \text{ pc}$ for the DM particles. The mass of the corresponding gas particles is of the order of $10^3 M_\odot$ (see table 1 in Wang et al. 2015 for more details about the resolution of the simulations).

The initial conditions for the zoom simulations were computed using a modified version of the GRAFIC2 code (Bertschinger 2011; Penzo et al. 2014). The level of refinement was chosen to keep a roughly constant numerical resolution across the probed mass range. The hydrodynamic simulations were run with the GASOLINE2 smoothed-particle-hydrodynamics (SPH) code by Wadsley, Stadel & Quinn (2004); Wadsley, Keller & Quinn (2017).

ESF-GASOLINE2 contains a subgrid model for turbulent mixing of metals and energy (Wadsley, Veeravalli & Couchman 2008). The calculation of heating and cooling includes the effects of a photoionizing UV background on metal-line cooling (Shen, Wadsley & Stinson 2010). Ritchie & Thomas's (2001) implementation of SPH allows a more accurate description of the multiphase structure of the gas and of the destruction of cold clumps. Limiting the time-step when it is much longer than the one for neighbouring particles (as proposed by Saitoh & Makino 2009) improves accuracy in the description of strong shocks. Finally, the number of neighbours used for SPH has been increased from 32 to 50.

Gas is available for star formation when it has temperature $T < T_{\text{th}} = 15000 \text{ K}$ and number density $n > n_{\text{th}} = 10.3 \text{ cm}^{-3}$. Gas that satisfies these conditions is converted into stars at the rate per unit volume:

$$\dot{\rho}_\star = \epsilon_{\text{sf}} \frac{\rho}{t_{\text{dyn}}} \quad (1)$$

where $\epsilon_{\text{sf}} = 0.1$ is the efficiency of star formation, ρ is the density of the gas available for star formation, and t_{dyn} is the dynamical time (Stinson et al. 2013 for technical details). Star formation is computed using time-steps of 0.8 Myr.

The NIHAO simulations contain two types of stellar feedback: pre-SN feedback (radiative feedback, winds from massive OB stars), also referred to as early stellar feedback (ESF), and SN blastwaves.

Assuming a Chabrier (2003) initial mass function and the relation of Torres, Andersen & Giménez (2010) between the mass and the luminosity of an individual star, a stellar population releases a luminous energy of $2 \times 10^{50} \text{ erg } M_\odot^{-1}$, most of which comes from short-lived massive stars. Pre-SN feedback is modelled assuming that a fraction ϵ_{ESF} of this energy is deposited as thermal energy in the interstellar medium (ISM), while cooling continues to be allowed

(Stinson et al. 2013). In the NIHAO simulations, $\epsilon_{\text{ESF}} = 0.13$ (Wang et al. 2015).²

SN feedback is implemented using the blastwave formalism described in Stinson et al. (2006), but without the superbubble feedback of Keller et al. (2014). The SN rate associated with each stellar particle is computed from its mass and age assuming an SN energy of 10^{51} erg for all stars between 8 and $40 M_{\odot}$. For a Chabrier (2003) initial stellar mass function, this assumption gives one SN every $120 M_{\odot}$ of stars. This energy is distributed to the surrounding gas. To prevent its immediate radiation by the artificially dense surrounding gas (the gas would not be so dense if the multiphase structure of the ISM were properly resolved), radiative cooling is switched off for ~ 30 Myr for some of the gas particles within the blastwave radius (Stinson et al. 2006 for details of how these particles are selected).

3 PHASES AND COMPONENTS

We classified gas particles into five categories: cold gas within the galaxy (the cold ISM), cold gas outside the galaxy but within the halo (the cold circumgalactic medium, CGM), hot gas within the galaxy (the hot ISM), hot gas outside the galaxy but within the halo (the hot CGM), and gas outside the halo (the intergalactic medium, IGM). This classification keeps into account both a particle’s position in physical space (inside the galactic radius, between the galactic radius and the virial radius, outside the virial radius; Section 3.1) and its location in the thermodynamic phase space (that is, on a density–temperature diagram; Section 3.2).

3.1 Virial radius and galactic radius

The first step of our analysis is to identify the centre, the virial quantities and the radius for each simulated galaxy.

We use the software PYNBODY (Pontzen et al. 2013) to identify the main halo in each NIHAO simulation. Its centre is determined by the AHF halo finder (Knollmann & Knebe 2009). The virial radius r_{vir} is the radius within which the mean density equals $\Delta_c \rho_c$, where ρ_c is the critical density of the Universe and Δ_c is the critical overdensity contrast computed with the fitting formulae by Bryan & Norman (1998), which give $\Delta_c \simeq 109$ at $z = 0$.

To separate the cold ISM from the cold CGM, we define the galactic radius r_g as the radius r_{HI} within which the surface density Σ_{gas} of the cold gas ($T < 15000$) is larger than the threshold for star formation in Kennicutt’s (1998) law, $\Sigma_{\text{th}} = 9 M_{\odot} \text{pc}^{-2}$.

The galactic radius r_g differs from r_{HI} only when r_{HI} decreases (r_g is not allowed to decrease) or when r_{HI} is smaller than the half-mass radius of the stars, $r_{\text{half-mass}}$, in which case we set $r_g = r_{\text{half-mass}}$, but that happens rarely and only in extremely gas poor objects. We also prevent the growth of r_g when r_{HI} increases by more than a factor of three from one output time-step to the next as this is invariably a sign of an ongoing merger. We impose these conditions to prevent the possibility that gas particles in the disc may be counted as expelled and reaccreted simply because of stochastic fluctuations in the value of r_{HI} .

Figs 1–3 are all for the same NIHAO galaxy g1.12e12, which exemplifies a typical massive spiral (it is the second most massive

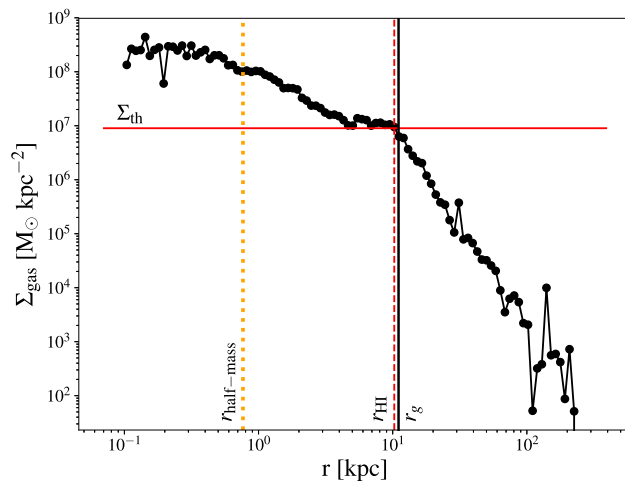


Figure 1. Gas surface density profile Σ_{gas} (black circles) and galactic radius r_g (vertical black solid line) at $z = 0$ for the NIHAO galaxy g1.12e12. Σ_{gas} is the mean surface density of gas with $T < 15000$ K. It is computed in circular annuli looking at the galaxy face on (i.e. from the direction of the angular momentum of the stars within r_{vir}). $\Sigma_{\text{th}} = 9 M_{\odot} \text{pc}^{-2}$ (horizontal red solid line) is the minimum surface density for star formation (Kennicutt 1998) and r_{HI} (vertical red dashed line) is the outermost radius for which $\Sigma_{\text{gas}} > \Sigma_{\text{th}}$ (the calculation is performed using a hundred radii equally spaced in $\log r$ between 0.1 kpc and r_{vir}). The vertical orange dotted line shows the half-mass radius of the stars, $r_{\text{half-mass}}$. We require that $r_g \geq \max(r_{\text{HI}}, r_{\text{half-mass}})$ and that the value of r_g should not decrease.

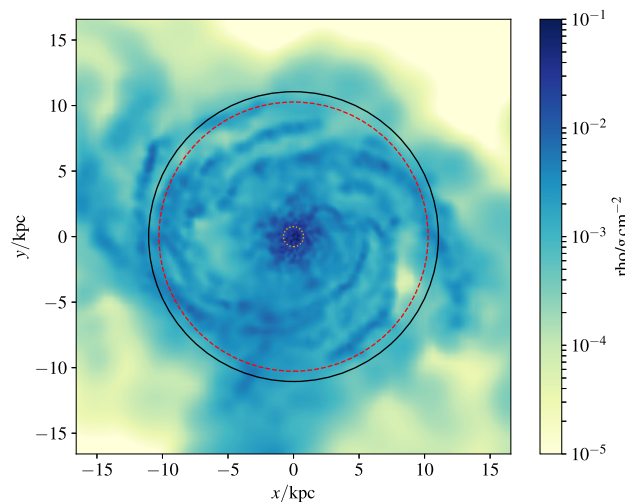


Figure 2. Cold-gas ($T < 15000$ K) distribution in a face-on image of the NIHAO galaxy g1.12e12. The black solid, the red dashed, and the orange dotted circles correspond to r_g , r_{HI} , and $r_{\text{half-mass}}$, respectively. The halo hosting this galaxy has $R_{\text{vir}} = 274$ kpc.

galaxy in the NIHAO series). We have given its identifier not only to be able to refer to it later on but also to enable comparison with other NIHAO publications. Fig. 1 illustrates the calculation of r_g graphically Fig. 2 shows the distribution of cold gas in g1.12e12 when the galaxy is viewed face-on; the red dashed circle corresponds to r_{HI} and shows that the outer edge of the HI disc has been correctly identified. Fig. 3 shows the distribution of cold gas (contours) and hot gas (colour scale) edge-on; one can see that the hot gas is strongly concentrated within the galaxy’s

²For comparison with previous work, the MaGICC simulations (Stinson et al. 2013) used a similar model with $T_{\text{th}} = 10^4$ K, $n_{\text{th}} = 9.3 \text{ cm}^{-3}$, $\epsilon_{\text{sf}} = 0.017$, and $\epsilon_{\text{ESF}} = 0.1$.

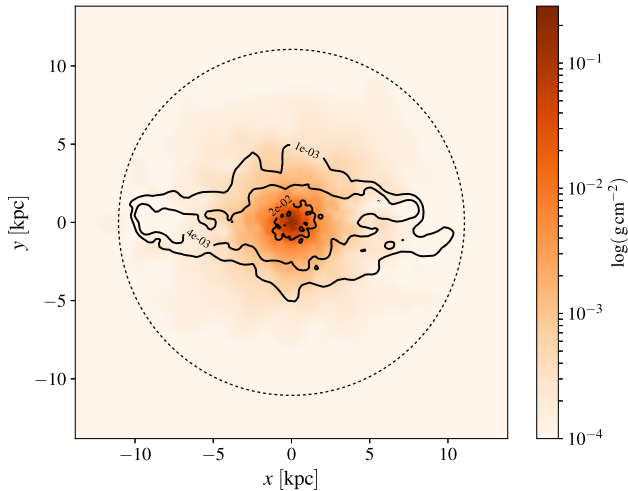


Figure 3. Spatial distribution of the cold and the hot ISM in the NIHAO galaxy g1.12e12. The galaxy is viewed edge-on. The dashed circle has radius r_g and is centred on the galactic centre. The colour scale shows the projected density distribution of the hot ISM (gas with $T > 15\,000$ K and $p > p_{\text{ISM}}$). The hot ISM is concentrated in the central part of the galaxy but the image also shows evidence for outflow. The contours show the projected density distribution of the cold ISM in units of g cm^{-2} . They follow the plane of the disc but they also show evidence that some cold gas has been lifted up by the hot ISM that vents out of the galaxy (see how the contours follow the hot ISM in the central region).

inner region, and that the cold gas is in a disc that thickens at the centre due to entrainment of cold gas by the outflowing hot component.

3.2 Gas phases

Fig. 4(a) shows the density–temperature distribution for all gas particles in g1.12e12. Figs 4(b) and (c) are identical to Fig. 4(a) except that they show only particles with $r < r_g$ and $r_g < r < r_{\text{vir}}$, respectively.

The gas along the white dashed line $T \propto \rho^{2/3}$ in the bottom left part of Fig. 4(a) belongs to the IGM because it is still outside r_{vir} (hence its absence in Figs 4b and c). This gas has such low density and temperature that it cannot radiate efficiently. Hence, its equation of state is virtually adiabatic.

When gas particles from the IGM reach the virial radius, they can take two paths (e.g. Kereš et al. 2005): they can stream on to the central galaxy in cold flows (cold accretion) or they can be shock-heated, form a hot low-density atmosphere (the hot CGM) and later cool on to the galaxy (Fig. 4a).

Galaxy g1.12e12 displays an atmosphere of shock-heated gas because it is a massive spiral with $M_{\text{vir}} = 1.29 \times 10^{12} M_{\odot}$. We chose a massive galaxy as an example precisely so that we would be able to see all thermodynamic phases on a density–temperature diagram. The hot CGM is virtually absent in haloes with $M_{\text{vir}} < 10^{11} M_{\odot}$, where all the baryons are accreted cold. However, the rest of our considerations from Fig. 4 do extend to lower masses.

Comparing Figs 4(b) and (c) shows that the hot high-density gas in the top right part of the density–temperature diagram and the hot low-density gas in top left part are different thermodynamic phases with different spatial distributions. The hot low-density gas is mainly at $r > r_g$. The hot dense gas is confined to $r < r_g$ and most of it is within the galaxy itself, although some of it is in a wind (Fig. 3). Hence, we refer to the hot low-density gas and

the hot dense gas as the hot CGM and the hot ISM, respectively.³ We emphasize that the latter is artificial and exists only because GASOLINE suppresses radiative cooling of gas particles heated by SN for ~ 30 Myr (Stinson et al. 2006). Hence, its only use is to identify gas that has been recently heated by SNe. Separating cold gas accreting on to a galaxy from cold gas in the galaxy is less straightforward but Figs 4(b) and (c) show that the star-forming gas is almost entirely within the galaxy.

Having identified the five phases that compose the gas in the NIHAO simulations (the IGM, the cold CGM, the cold ISM, the hot ISM, and the hot CGM), we are now ready to set out precise criteria to assign a phase to each particle.

An obvious choice is to use a temperature criterion (the maximum temperature for star formation) to separate cold gas from hot gas. All the hot CGM (Fig. 4c) and the hot ISM (Fig. 4b) are above the lines AB and DE, corresponding to $T = 15\,000$ K. The problem of using a temperature criterion alone is that this choice would attribute part of the cold gas to the hot CGM, since cold streams within haloes are in temperature range $T \sim 10\,000 - 20\,000$ K.

To separate the cold filaments from the hot atmosphere, we observe that the adiabat BC corresponds to the entropy of IGM. Cooling pushes gas to the right of this adiabat. Shock-heating pushes it to the left. Hence, it is natural to use this adiabat at $T > 15\,000$ K to separate the cold CGM from the hot CGM.

An adiabat is specified by its temperature at a given density. Let this density be the mean baryonic density of the Universe $\bar{\rho}_b = \Omega_b \rho_c \simeq 6 M_{\odot} \text{kpc}^{-3}$. The adiabat that fits best the temperature–density relation around $\bar{\rho}_b$ is:

$$T = 4200 \text{ K} \times 10^{-0.25a(z) - 0.86a^2(z)} \left(\frac{\rho}{\bar{\rho}_b} \right)^{2/3}, \quad (2)$$

where z is redshift and $a(z) = z/(1+z)$.

To separate cold gas from gas that is sufficiently hot to vent out of the galaxy, we use a pressure criterion. The line CD is an isobar with pressure equal to the pressure $p_{\text{ISM}} \simeq 2 \times 10^{-13} \text{ dyn cm}^{-2}$ of the cold dense star-forming ISM (i.e. the molecular gas). Gas with higher temperature and thus higher pressure will leave this phase but will remain within the galaxy as cold neutral gas if p_{ISM} is not exceeded by a large factor. Hence, gas with $T \leq 15\,000$ K is still considered part of the cold ISM even if it has pressure $p > p_{\text{ISM}}$. The dotted-dashed line in Fig. 4(a) shows that this choice is equivalent to separating the cold ISM from the hot ISM at $p = 100 p_{\text{ISM}}$.

We note that the ISM in the NIHAO simulations is very different from the one of McKee & Ostriker (1977), based on the assumptions of different phases in pressure equilibrium. The hot ISM is not in pressure equilibrium with the cold ISM. It is continually venting out of galaxies (and cooling back on to them, as we shall see in Section 9).

The distinction between cold gas in the galaxy and cold gas in the halo is the only one that cannot be performed based on the thermodynamic phase-space diagram alone and relies on our measurement of r_g .

Particles at the borders between phases may complicate our analysis because small fluctuations in density and temperature may be misinterpreted as ejection and reaccretion. For this reason, a particle is moved from one phase to another only when the threshold value has been exceeded by 10 per cent. A particle initially outside

³The hot ISM is identified on the basis of a purely thermodynamic criterion. Hence, it includes both the genuine hot ISM and a hot dense wind component (the gas in the top right corner of Fig. 4c).

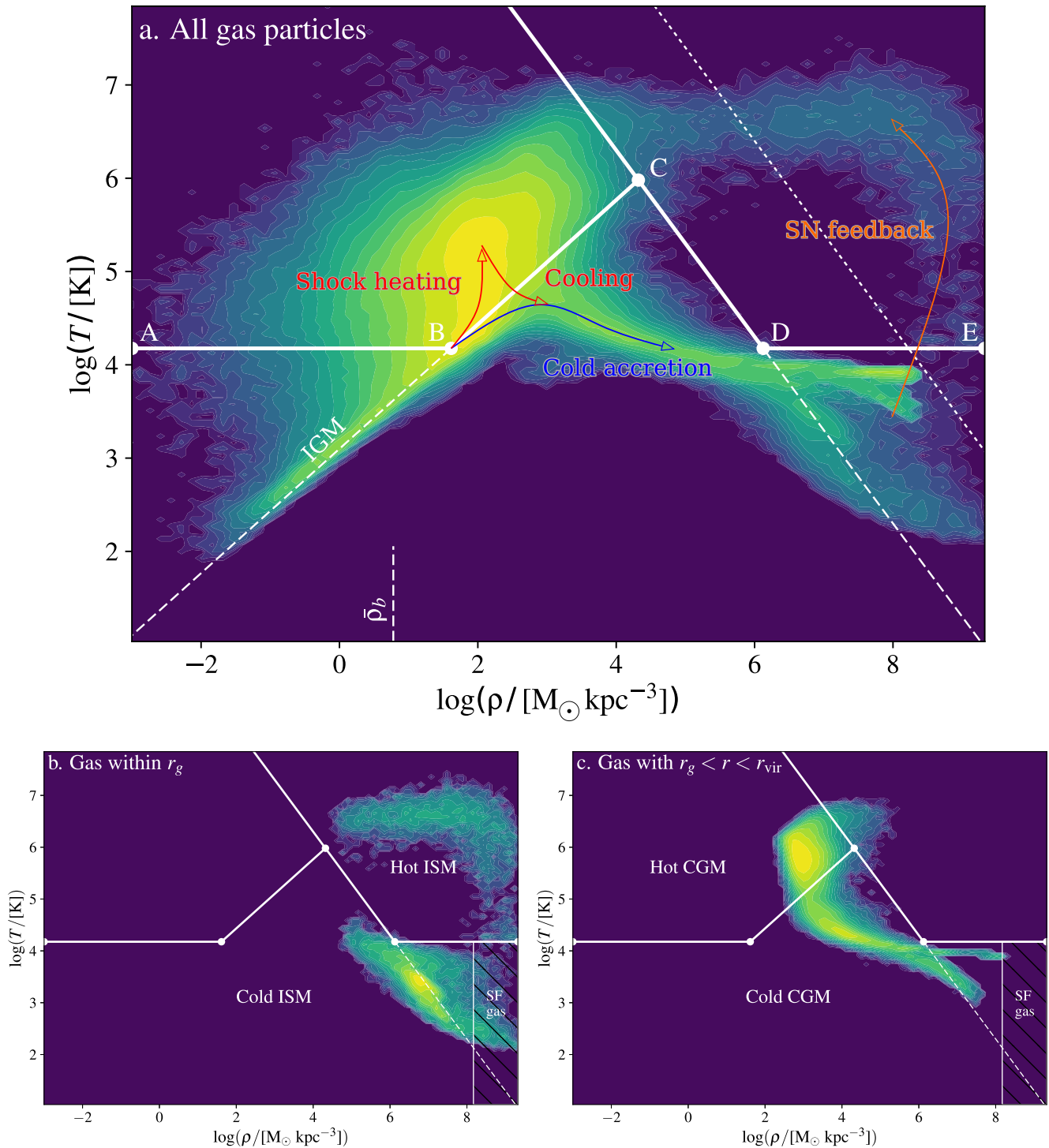


Figure 4. Density–temperature diagram for g1.12e12 ($M_{\text{vir}} = 1.12 \times 10^{12} M_{\odot}$). The three panels refer to: (a) all the gas in the simulated region, (b) gas within the galaxy ($r \leq r_g$), and (c) gas within the halo but outside the galaxy ($r_g < r < r_{\text{vir}}$). The white solid lines separate the different thermodynamic phases and are the same in all three panels. AB and DE correspond to the isotherm $T = 15000$ K. BC is an adiabat with entropy equal to the one of the IGM (equation 2). The isobar CD and the white dotted-dashed line in panel (a) correspond to $p = p_{\text{ISM}} \simeq 2 \times 10^{-13}$ dyn cm $^{-2}$ and $p \simeq 2 \times 10^{-11}$ dyn cm $^{-2}$, respectively. The hatched rectangles in the bottom right corners of panels (b) and (c) highlight the cold dense star-forming gas. They show that this gas is almost entirely within the galaxy. The vertical white dashed line in panel (a) shows the mean baryonic density of the Universe, $\bar{\rho}_b$.

the galaxy moves inside the galaxy when it moves to $r < 0.9r_g$. A particle within the galaxy is counted as ejected only when it moves to $r > 1.1r_g$. A particle crosses the lines AB or DE and moves to the hot CGM or ISM only when its temperature raises

above 16 500 K. A particle crosses the lines AB or DE and moves to the cold phase only when its temperature drops below 13 500 K. The same 10 per cent elastic border is applied to the pressure p_{ISM} that separates the cold gas from the hot ISM (line CD). The only

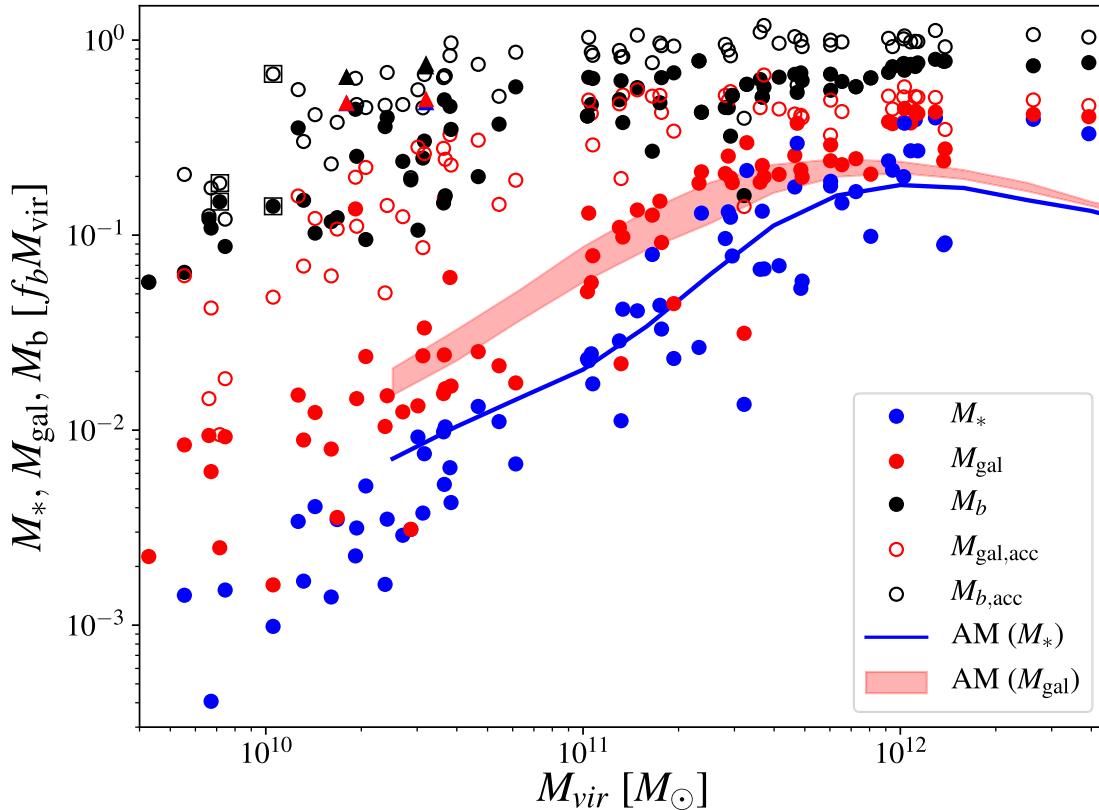


Figure 5. Stellar masses (filled blue circles), galaxy masses (stars plus cold galactic gas; filled red circles), and total baryonic masses (filled black circles) within r_{vir} at $z = 0$ as a function of M_{vir} . All masses are expressed in units of $f_b M_{\text{vir}}$. $M_{\text{gal,acc}}$ ($M_{\text{b,acc}}$) is the total mass of the baryons particles identified as belonging to the galaxy (the halo) at any output time-steps (open red and black circles, respectively). The filled triangles are the same as the filled circles for a couple of galaxies that have been resimulated without stellar feedback. The circles enclosed in squares are a couple of objects that we have highlighted to compare their behaviours. The blue curves show the stellar mass–halo mass relation from the AM studies of Papastergis et al. (2012; solid curve). The red shaded area shows the galaxy mass–halo mass relation from the AM study of Papastergis et al. (2012), who computed the baryonic mass function of galaxies using both optical and HI data.

exception is the adiabat BC, to which we apply a tolerance of 5 per cent only, since many particles in the hot CGM have entropies just above the critical entropy that separates the hot CGM from the cold CGM.

4 BARYON CONTENT OF GALAXIES AND HALOES

Having introduced the NIHAO simulations (Section 2) and our method to analyse them (Section 3), we are now ready to present our results. We start from the baryon fraction within galaxies and haloes (current section). Later we shall investigate the astrophysical processes that are responsible for the baryon fractions we find.

Fig. 5 shows the stellar mass M_* (filled blue circles), the baryonic mass M_{gal} within the galaxy (stars and cold gas, filled red circles), and the total baryonic mass M_b within r_{vir} (filled black circles) for the NIHAO galaxies at $z = 0$. All masses are shown as a function of M_{vir} and in units of $f_b M_{\text{vir}}$, where $f_b = 0.16$ is the cosmic baryon fraction.

The filled blue circles (our $M_* - M_{\text{vir}}$ relation) are in reasonable agreement with the results from abundance matching (AM) by Papastergis et al. (2012; blue solid curve). The only small tension with the observations is that our ultradwarves tend to be of lower mass than in AM studies (but notice that our procedure to estimate r_g lowers M_* in gas-poor galaxies, and that passive ultra-

dwarves fall into this category). Throughout this article, dwarves are galaxies with $3 \times 10^7 M_\odot \lesssim M_* \lesssim 3 \times 10^9 M_\odot$ (corresponding to $3 \times 10^{10} M_\odot \lesssim M_{\text{vir}} \lesssim 3 \times 10^{11} M_\odot$); ultradwarves and normal spirals are galaxies below and above this range of masses, respectively.

The filled red circles show the same relation for the baryonic galaxy mass M_{gal} (the mass of stars and the cold ISM). Comparing them to the AM relation by Papastergis et al. (2012; red shaded area), who measured $M_{\text{gal}} = M_* + 1.4M_{\text{HI}}$ from a sample with HI data for each galaxy (the factor 1.4 accounts for helium) shows that the agreement with the total galaxy mass is also satisfactory. The simulations’ ability to reproduce the stellar and baryonic masses of galaxies simultaneously for a same data set proves that gas fractions are reproduced reasonably well, at least for $M_{\text{vir}} \gtrsim 3 \times 10^{10} M_\odot$.

The filled black circles show that the halo baryon content is 60 – 80 per cent of the cosmic baryon fraction for normal spirals with $M_{\text{vir}} > 3 \times 10^{11} M_\odot$ but drops to 10 – 20 per cent of the cosmic baryon fraction for ultradwarves with $M_{\text{vir}} < 2 \times 10^{10} M_\odot$. As already reported by Wang et al. (2017), these findings are consistent with those by Brook et al. (2014), Hopkins et al. (2014), and Mitchell et al. (2018), even though our baryon fractions tend to be slightly lower.

The open red circles replace M_{gal} with $M_{\text{gal,acc}}$, the sum of the masses of all particles ever identified as belonging to the galaxy. $M_{\text{gal,acc}}$ includes M_{gal} , but it also includes the masses of

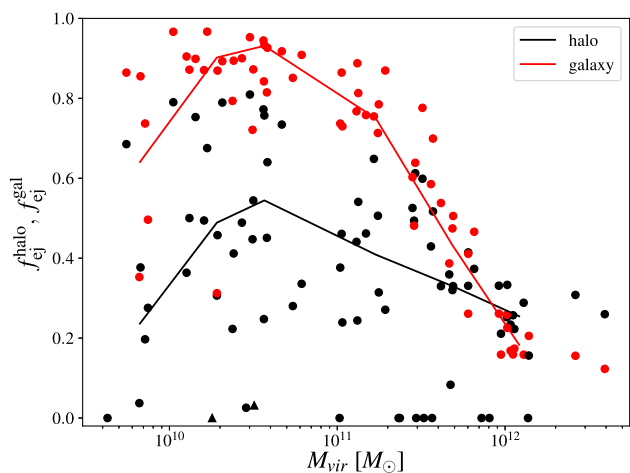


Figure 6. Ejected mass fraction from the galaxy (red) and the halo (black). The symbols correspond to individual galaxies. The curves are medians in bins of halo mass. Ejected fractions are relative to the total baryonic masses accreted by galaxies or haloes over the lifetime of the Universe. They have been computed counting all baryons that have been inside the galaxy or the halo at some point but are no longer there at $z = 0$. The triangles are two objects that have been resimulated without SN feedback. Both are consistent with $f_{\text{ej}}^{\text{halo}} \simeq 0$.

all the particles that have been expelled from the galaxy. When $M_{\text{gal,acc}}/M_{\text{gal}} \sim 10$, that means that 90 per cent of the baryons accreted by the galaxy have been ejected from the galaxy.

The impact of the expulsion of matter on the baryon content of galaxies and haloes can be quantified more precisely by introducing the ejected fractions:

$$f_{\text{ej}}^{\text{gal}} = 1 - \frac{M_{\text{gal}}}{M_{\text{gal,acc}}}, \quad (3)$$

$$f_{\text{ej}}^{\text{halo}} = 1 - \frac{M_{\text{b}}}{M_{\text{b,acc}}}. \quad (4)$$

Fig. 6 (red curve) shows that $f_{\text{ej}}^{\text{gal}}$ decreases from about 95 per cent at $M_{\text{vir}} \sim 4 \times 10^{10} M_{\odot}$ to about 20 per cent at $M_{\text{vir}} \sim 10^{12} M_{\odot}$. Low $f_{\text{ej}}^{\text{gal}}$ are also measured in ultradwarves where there has not been enough star formation to produce massive outflows.

In dwarves, the median ejected fraction from the halo is about half the ejected fraction from the galaxy (one-third in the case of ultradwarves). The difference between $f_{\text{ej}}^{\text{halo}}$ and $f_{\text{ej}}^{\text{gal}}$ becomes smaller at higher masses. At $M_{\text{vir}} > 8 \times 10^{11} M_{\odot}$, nearly all galaxies have $f_{\text{ej}}^{\text{halo}} > f_{\text{ej}}^{\text{gal}}$ (Fig. 6, black and red circles). This finding is only apparently paradoxical because *most of the baryons ejected from haloes do not come from galaxies*. They come from halo gas that has been pushed out by galactic winds or unbound by the thermal energy that SNe have deposited into it (Fig. 7). The NIHAO sample selects galaxies that are fairly isolated (Section 2), but some of the gas that leaves the halo without having ever been part of the galaxy may also come from satellites that have come in and out of the virial sphere.

One may wonder to what extent the baryons that are ejected from haloes but do not come from galaxies may be in fact winds coming from small satellites. We have verified that such winds contribute to the outflow rate from the halo but that our qualitative results do not change if we remove DM substructures and the associate baryons from our analysis.

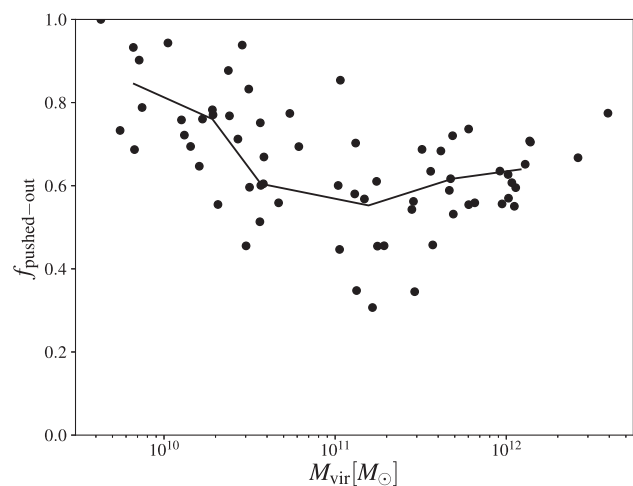


Figure 7. Baryon fraction that has never been inside the galaxy relative to the total mass of the gas particles expelled from the halo (symbols: individual galaxies and curve: median in bins of halo mass). For $M_{\text{vir}} \lesssim 3 \times 10^{10} M_{\odot}$, $f_{\text{pushed-out}}$ is close to unity (median value > 0.8). Hence, most of the gas ejected from the halo comes from the CGM and not from the galaxy. The larger M_{vir} , the more difficult for winds to entrain halo gas ($f_{\text{pushed-out}}$ decreases). The trend is reversed at high masses because dilute shock-heated gas is more easily pushed out than dense cold gas.

The existence of haloes with $M_{\text{b}}/(f_{\text{b}}M_{\text{vir}}) \sim 0.06$ (Fig. 5) while $f_{\text{ej}}^{\text{halo}} \lesssim 0.8$ for all objects (Fig. 6) proves that ejective feedback is not the only mechanism that is responsible for the low baryon fractions of many NIHAO haloes.

To demonstrate that two different processes are at work, we have identified two haloes with $M_{\text{b}}/(f_{\text{b}}M_{\text{vir}}) \sim 0.1$. They correspond to the circles enclosed in squares in Fig. 5. The halo with $M_{\text{vir}} \sim 1.5 \times 10^{10} M_{\odot}$ has accreted 60–70 per cent of the universal baryon fraction but has lost ~ 70 per cent of its baryons. The other halo (the one with $M_{\text{vir}} \sim 8 \times 10^9 M_{\odot}$) has accreted less than 20 per cent of the universal baryon fraction but has retained most of the accreted baryons. Ejective feedback has played a major role in the first halo but not in the second one, where baryons never made it into the virial radius in the first place. The presence of many haloes with $M_{\text{b,acc}} \ll f_{\text{b}}M_{\text{vir}}$ (Fig. 5, open circles) demonstrates that this is not an isolated case and that inability to accrete is a significant limiting factor for the growth of galaxies in haloes with $M_{\text{vir}} < 3 \times 10^{10} M_{\odot}$ (Nelson et al. 2015; Wang et al. 2017; Mitchell et al. 2018).

To understand whether this inability is due to stellar feedback or heating by a photoionizing UV background, we have resimulated two galaxies without stellar feedback (triangles in Fig. 5). Fig. 8 shows the flow lines and the temperature distribution of gas in the less massive of the two galaxies ($M_{\text{vir}} = 1.6 \times 10^{10} M_{\odot}$) with and without stellar feedback (left- and right-hand panels, respectively). The right-hand panel shows that re-ionization feedback alone (the only feedback that is present when SN feedback and early stellar feedback are turned off) has limited effect at $M_{\text{vir}} \gtrsim 2 \times 10^{10} M_{\odot}$. Hence, gas streams into the virial radius (dashed circle), accretes on to the galaxy (Fig. 8, right) and forms stars so efficiently that, in Fig. 5, the blue triangles (M_{\star}) hide the red triangles (M_{gal}) almost entirely. The left-hand panel of Fig. 8 shows that including stellar feedback changes the dynamics of the gas completely. Stellar-driven outflows terminate with a shock front at $r = 6r_{\text{vir}}$ (dashed circle) and prevent the infalling gas from coming in. The only gas that accretes on to the galaxy is warm ($10^4 \text{ K} < T < 10^5 \text{ K}$) gas that cools within

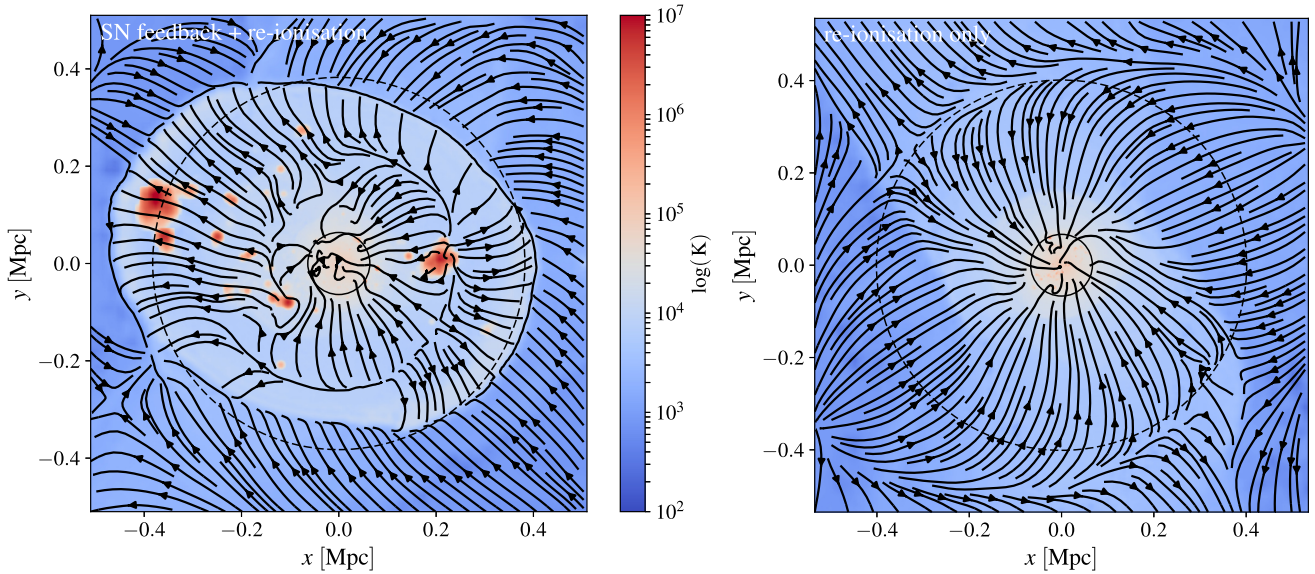


Figure 8. Temperature map and flow lines for the gas in the NIHAO simulation g1.5e10 ($M_{\text{vir}} = 1.6 \times 10^{10} M_{\odot}$ at $z = 0$), which corresponds to the lower mass triangle in Fig. 5. The left-hand panel shows the standard NIHAO simulation inclusive of both stellar feedback and feedback from a photoionizing UV background. The right-hand panel shows the same system re-simulated with feedback from a photoionizing UV background only. The small circle at the centre shows the virial radius. The large dashed circle corresponds to $r = 6r_{\text{vir}}$. Gas particles shown in these figures are part of the zoom-in simulated region, which extends up to ~ 1 Mpc for this galaxy.

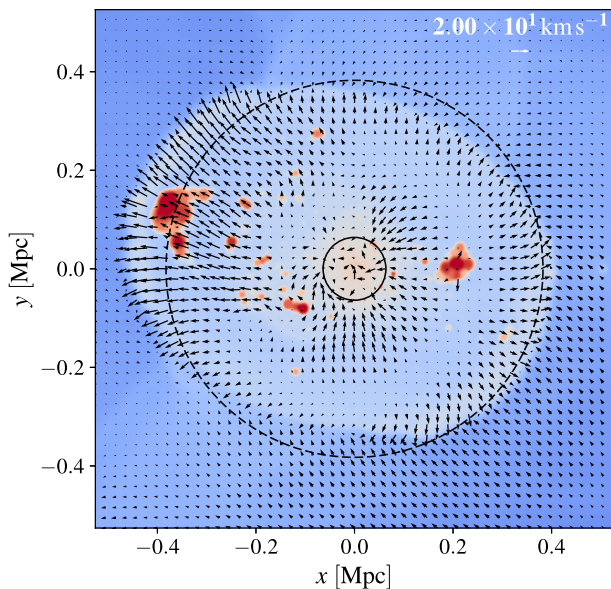


Figure 9. Temperature map and velocity for the gas in the NIHAO simulation g1.5e10 (with stellar feedback and UV background), Fig. 8 shows the corresponding flow lines. The small circle at the centre shows the virial radius. The dashed circle corresponds to $r = 6r_{\text{vir}}$. Colour coding shows temperature and is the same as in Fig. 8. The outflow speed is much higher at 10 o'clock (20 km s^{-1} at $6r_{\text{vir}}$) than in the other directions. This is also the direction along which the hot bubbles at $T \sim 10^7$ K are coming out.

a sphere of radius $4r_{\text{vir}}$. Fig. 9 shows that the hot outflowing gas drives the expansion of a shock front that propagates in the wind's direction.

In conclusion, without stellar feedback, the NIHAO galaxies have $M_{\star} > 0.4f_{\text{b}}M_{\text{vir}}$ down to $M_{\text{vir}} \sim 2 \times 10^{10} M_{\odot}$ (Fig. 5, blue triangles). SNe reduces M_{\star} by two orders of magnitude but not through ejective feedback alone. One order of magnitude is gained

at a much lower energy cost by preventing gas from coming while it is still far away. This pre-emptive feedback is dominant mechanism in ultradwarves with $M_{\text{vir}} \lesssim 10^{10} M_{\odot}$.

5 STAR FORMATION

The efficiency of star formation $\epsilon_{\text{sf}} = 0.1$ adopted in the NIHAO simulations (Section 2) is calibrated on the local Kennicutt (1998) law and we have already seen that this assumption leads to integrated star formation rates (SFRs) in reasonable agreement with the observed stellar masses of galaxies (Fig. 5). Here, however, we want to check the NIHAO galaxies' star formation efficiencies directly by comparing the global star formation time-scale

$$t_{\text{sf}} = \frac{M_{\text{gas}}}{\text{SFR}} \quad (5)$$

with observations from the Herschel Reference Survey (HRS, Boselli et al. 2014).

In Fig. 10, the red circles show the $t_{\text{sf}}-M_{\star}$ relation in the NIHAO simulations when M_{gas} (equation 5) is the total mass of the cold ISM. The red shaded area is a linear fit to the data for HI-and-CO-detected late-type gas-rich galaxies (Boselli et al. 2014). In these observations, $M_{\text{gas}} = 1.3(M_{\text{HI}} + M_{\text{H}_2})$, where the factor 1.3 accounts for the presence of helium. The NIHAO simulations are in reasonable agreement with the data and recover the observational scaling relation $t_{\text{sf}} \propto M_{\star}^{-0.7}$.

For $M_{\star} < 10^9 M_{\odot}$, t_{sf} becomes longer than the age of the Universe. Hence, inefficient accretion and the expulsion of gas from gas galaxies and haloes are not the only mechanisms that slow down the conversion of baryons into stars. Inefficient star formation plays a role, too.

The fraction of the available gas that can be converted into stars in a time t is $f = 1 - \exp(-t/t_{\text{sf}})$. When t_{sf} is much shorter than the age of the Universe, f is of order unity, but that is not true in galaxies with $M_{\star} \sim 10^8 M_{\odot}$, where t_{sf} can take all values between 10 and 200 Gyr (Fig. 10, red circles). Taking a median value $t_{\text{sf}} \sim 40$ Gyr

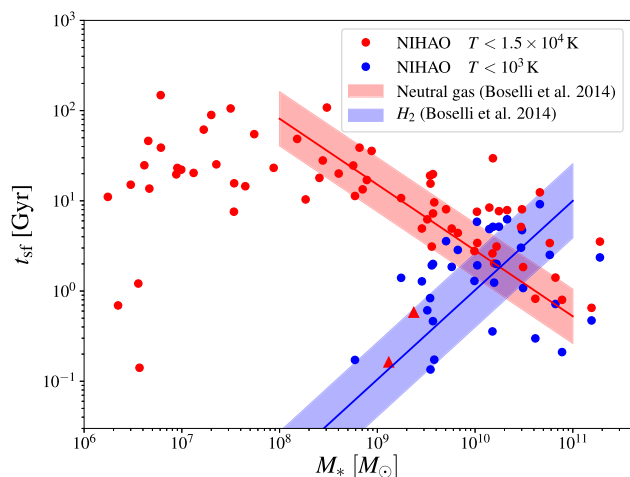


Figure 10. The relation between star formation time-scale $t_{\text{sf}} = M_{\text{gas}}/\text{SFR}$ and stellar mass M_* in the NIHAO simulations (symbols) and the HRS (Boselli et al. 2014; shaded areas). Red and blue refer to the total mass of cold neutral gas within a galaxy and to the mass of dense cold molecular mass, respectively. In the NIHAO simulations, this gas is identified by using the condition $T < 1000$ K. In Boselli et al. (2014), the H_2 mass is measured from CO data assuming a CO/H_2 ratio that depends on luminosity (Boselli, Lequeux & Gavazzi 2002). The red triangles show the results for the simulations without SN feedback. In those objects nearly all the gas has $T < 1000$ K.

and assuming $t \sim 10$ Gyr limits the mass that can form stars to ~ 20 per cent of the baryons available in the cold ISM.

The inefficiency of star formation at low M_* derives from the cold ISM’s inability to condense into molecular clouds: we have inspected the density–temperature diagram for low-mass galaxies and we have verified that dwarves with $t_{\text{sf}} \gtrsim 10$ Gyr contain very little gas with $T < 1000$ K. We refer to this phenomenon with the name of anticondensation feedback.

The same conclusion can also be reached from Fig. 10. The star formation time-scale for gas with $T < 1000$ is always short and decreases at low M_* (in agreement with Boselli et al. 2014). Hence, the molecular gas always forms stars effectively.

The proof that anticondensation feedback is a genuine feedback effect is that nearly all the gas is converted into stars when stellar feedback is switched off (in Fig. 5, the blue triangles overlap with the red ones).

The data of Boselli et al. (2014) do not extend below $M_* = 10^8 M_\odot$. In fact, they contain very few galaxies with $M_* < 10^9 M_\odot$. However, the NIHAO simulations predict that the star formation time-scale for the HI gas cannot keep rising like $t_{\text{sf}} \propto M_*^{-0.7}$ at low masses: dwarves galaxies with $M_* < 10^8 M_\odot$ form so few stars that there are not enough SNe for effective feedback. Ultradwarves with $M_* < 10^7 M_\odot$ are predicted to exhibit highly stochastic SFRs with intense starbursts followed by long periods of quiescence. Fig. 10 shows that in these systems t_{sf} can take all values between 0.3 and 200 Gyr.

6 MASS LOADING

In Section 4, we have seen that 90 per cent of the baryons that accrete on to a dwarf galaxy ($M_* \sim 10^8 M_\odot$) are expelled from the galaxy and that this fraction goes down to ~ 20 per cent for massive spirals such as Andromeda and the Milky Way. We have also seen that this is not the only reason why dwarf galaxies have very low

M_*/M_{vir} mass ratios. Inability to accrete (Section 4) and inefficient star formation (Section 5) play a role, too. We now want to take our analysis one step forward and establish a direct link between star formation and outflows. Our goal is to extract predictions that may be compared to observations and also to derive physically motivated prescriptions that may improve the description of feedback in SAMs of galaxy formation.

Let us start by introducing the mass-loading factor:

$$\eta = \frac{1}{\text{SFR}} \frac{\Delta M_{\text{out}}}{\Delta t}, \quad (6)$$

where Δt is the time between the consecutive output times t and $t + \Delta t$, and where ΔM_{out} is the mass of the gas particles ejected between t and $t + \Delta t$.

In this article, we consider four different definitions of ejected particles:

- (i) particles that are in the cold ISM at t but no longer belong to it at $t + \Delta t$;
- (ii) particles that are in the cold ISM at t but have moved to $r > r_g$ at $t + \Delta t$;
- (iii) particles that are in the cold ISM at t but have moved or are in the process of moving to the cold CGM at $t + \Delta t$;
- (iv) particles that are within the virial radius at t but have moved to $r > r_{\text{vir}}$ at $t + \Delta t$.

Each definition serves a different purpose. The first gives the rate at which gas ceases to be available for star formation. The second gives the rate at which gas flows out of the galaxy. The third is useful to compare with observations that trace the cold molecular component.⁴ The fourth gives the rate at which gas leaves the halo. We use $\eta_{\text{cold ISM}}$, η_{gal} , $\eta_{\text{gal}}^{\text{cold}}$, and η_{halo} , respectively, to denote the mass-loading factors computed with these four definitions. Our definitions are such that, for each individual galaxy, $\eta_{\text{gal}}^{\text{cold}} \leq \eta_{\text{gal}} \leq \eta_{\text{cold ISM}}$ by construction.

Fig. 11 compares the mass-loading factors from the NIHAO simulations with those estimated by Cicone et al. (2014) for M82, NGC 253, NGC 2146, NGC 3256, and NGC 3628 (we took the sample by Cicone et al. and removed all objects in which they found evidence for an active galactic nucleus). The greatest difficulty of this comparison was to find reliable masses for such nearby objects. Their apparent sizes are so large that the SDSS and 2MASS miss a lot of the light from the outer regions. We therefore reverted to extinction-corrected B -, V -, and R -band photometry from the Third Reference Catalogue of Bright Galaxies (de Vaucouleurs et al. 1991) and converted luminosities into masses using stellar mass-to-light ratios from PEGASE (Fioc & Rocca-Volmerange 1997), which kept colours into account. The good agreement of η_{gal} (yellow circles) with the observations becomes excellent when the comparison is based on the cold component $\eta_{\text{gal}}^{\text{cold}}$ (filled blue circles) only, as it should be, since the points by Cicone et al. (2014) are based on CO data. We stress that no free parameter was tuned to obtain this result.

The point from Falgarone et al. (2017) refers to a different redshift range ($2 < z < 2.6$ instead of the local Universe) and carries considerable uncertainty.⁵ We have added it nevertheless because our $\eta_{\text{gal}}^{\text{cold}} - M_*$ relation at $z = 2.3$ is not significantly different from the one at

⁴Definition (iii) includes hot ISM’s particles in transit from the cold ISM to the cold CGM but few particles take this route. When they do, the time spent in the intermediate hot-ISM phase is very short. Therefore, particles identified with definition (iii) are effectively cold at all times.

⁵Falgarone et al. (2017) used ALMA data to measure mass-loading factors for six submillimetre galaxies with $\text{SFR} = 300 - 13000 M_\odot \text{ yr}^{-1}$ at $2 < z <$

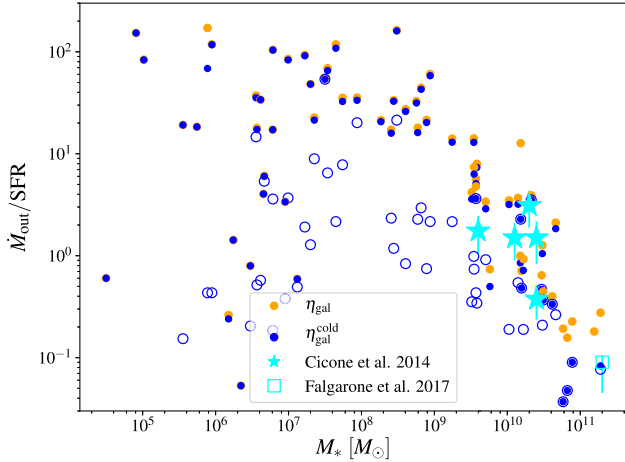


Figure 11. Mass-loading factors for the total outflow rate (η_{gal} , yellow circles) and the cold component only ($\eta_{\text{gal}}^{\text{cold}}$, blue filled circles) versus stellar mass M_* in the NIHAO simulations. The blue open circles show $\eta_{\text{gal}}^{\text{cold}}$ when we consider only gas with $T < 10^3$ K. The stars with error bars are the data for five nearby galaxies (M 82, NCG 3256, NGC 3628, NGC 253, and NGC 2146). They are based on CO observations of molecular outflows (Cicone et al. 2014). The open cyan square is obtained by stacking six submillimetre galaxies at $z \sim 2.3$ (Falgarone et al. 2017). The NIHAO simulations predict mass-loading factors that decrease with M_* at $M_* > 10^8 M_\odot$ and stochastic mass-loading factors at lower masses.

$z = 0$ and this point is our only constraint for $\eta_{\text{gal}}^{\text{cold}}$ at high masses. Falgarone et al.’s (2017) data point shows that the steep decrease of $\eta_{\text{gal}}^{\text{cold}}$ with M_* is real and that the NIHAO simulations reproduce the correct relation over two orders of magnitude in stellar mass.

The mass-loading factor $\eta_{\text{gal}}^{\text{cold}}$ depends on our definition of cold gas. The filled blue circles in Fig. 11 consider all the gas identified as belonging to the cold ISM in Fig. 4(b). As the mass-loading factor from Cicone et al. (2014) and Falgarone et al. (2017) are inferred from Co and CH + tracers, respectively one may wish to apply a more restrictive criterion and consider cold only the molecular gas ($T < 1000$ K). When we do it the blue filled circles become the blue open circles, which are in good agreement with the observations at high masses but seems to underestimate mass-loading factor at lower masses. One should notice, however, that only massive NIHAO galaxies exhibit a dense molecular phase distinct from the atomic gas. In lower mass galaxies, the cold ISM does not contain enough particles for this phase to be properly resolved.

One should also be aware that Cicone et al. measured outflow rates by assuming that the outflowing clouds populate the spherical (or multiconical) region affected by outflows uniformly. This assumption is reasonable if outflows are continually refilled with clouds ejected from the disc but it overestimates $\eta_{\text{gal}}^{\text{cold}}$ for stochastic outflows (which we find to be prevalent in dwarf galaxies but not in massive objects, see Section 7).

The comparison with observations has made us confident that the NIHAO simulations approximate the Universe reasonably well. With this confidence, we are now ready to study the dependence of

2.6 and found values of ~ 0.1 for all of them (within the large uncertainties of their measurements). We have no information about the masses of these galaxies but we have estimated their order of magnitude from the SFR – M_* at $z \sim 2$ (Daddi et al. 2007; Tasca et al. 2015). $M_* \sim 2 \times 10^{11} M_\odot$ is the average stellar mass of a galaxy with SFR = $300 M_\odot \text{ yr}^{-1}$ at $z = 2$. This calculation is more likely to overestimate M_* than underestimate it.

η on the virial velocity v_{vir} and the virial mass M_{vir} for each of our four definitions of the mass-loading factor. The red, yellow, blue, and black symbols in Fig. 12(a) correspond to $\eta_{\text{cold ISM}}$, η_{gal} , $\eta_{\text{gal}}^{\text{cold}}$, and η_{halo} , respectively, while the red, yellow, blue, and black solid lines are log–log linear least-squares fits of the form:

$$\eta = \eta_0 \left(\frac{v_{\text{vir}}}{100 \text{ km s}^{-1}} \right)^\alpha, \quad (7)$$

to the symbols of matching colours.

Table 1 gives the best-fitting parameters η_0 and α for each definition of η when we stop the fits at $v_{\text{vir}} = 45 \text{ km s}^{-1}$. Below this virial velocity, the scatter in η becomes huge and the fits would be meaningless. This scatter is due to the stochastic SFR histories of dwarves, characterized by intermittent bursts of star formation interspersed between periods of quiescence (there can be no outflows without type II SNe and there can be no type II SNe without star formation).

The yellow and the blue lines in Fig. 12(a) have been computed fitting all the points at $v_{\text{vir}} > 45 \text{ km s}^{-1}$ but we have stopped them at $v_{\text{vir}} = 67 \text{ km s}^{-1}$ because it would be meaningless to extrapolate the fits where $\eta_{\text{gal}}^{\text{cold}} > \eta_{\text{gal}}$ (the cold component cannot larger than the total outflow rate) or where $\eta_{\text{cold ISM}} < \eta_{\text{gal}}$ (including gas that is heated but remains within the galaxy cannot give lower outflow rates than excluding it). The steeper decline of the blue line compared to the yellow line and of the yellow line compared to the red line shows that at low masses most of the mass is the cold component, while in massive galaxies plenty of gas is heated by SNe but does not escape.

The cyan points with error bars in Fig. 12(a) compare the results of the NIHAO simulations to measurements from MGII absorption lines by Bouché et al. (2012, triangles), Kacprzak et al. (2014, square), and Schroetter et al. (2015, circles), which trace warm ionized gas with outflow speeds of a few hundred kilometres per seconds which are intermediate between those of cold molecular gas and hot X-ray gas. We also remark that the red line in Fig. 12(b) is consistent with the η_c – M_{vir} relation at $z = 0$ in Anglés-Alcázar et al. (2017) and that the black line is consistent with the η_{loss} – M_{vir} relation at $z = 0$ in the same article. Anglés-Alcázar et al. (2017) define η_c as the ratio of the total mass ejected in winds from the ISM (regardless of its fate after ejection) to the total stellar mass formed *in situ* from early times down to $z = 0$, and η_{loss} as the same ratio counting only the particles that have been ejected and never reaccreted.

Besides the sporadic star formation of dwarves discussed above (also see Section 5), there is another physical reason why η cannot not keep growing as v_{vir}^α with $\alpha \ll -2$ at low v_{vir} (Cattaneo et al. 2017). Outflows require speeds greater than the escape speed v_{esc} and this sets an upper limit to outflow rate:

$$\dot{M}_{\text{out, max}} = \frac{2\epsilon_{\text{SN}} E_{\text{SN}} \Psi_{\text{SN}}}{v_{\text{esc}}^2} \text{SFR}, \quad (8)$$

(Silk 2003), where Ψ_{SN} is the number of SNe per unit stellar, E_{SN} is the energy released by one SN, and ϵ_{SN} is the efficiency with which the energy from SNe is converted into kinetic energy. As $v_{\text{esc}} \propto v_{\text{vir}}$, equation (8) implies $\eta_{\text{max}} = \dot{M}_{\text{out, max}}/\text{SFR} \propto v_{\text{vir}}^{-2}$ if we assume that ϵ_{SN} is constant.

In reality, there is no reason why ϵ_{SN} should be constant. On the contrary, there are physical reasons why ϵ_{SN} should decrease with v_{vir} . One of them is that massive haloes contain massive atmospheres of shock-heated gas (Dekel & Birnboim 2006). The efficiency of SNe is limited by the radiative efficiency of the heated gas, which is higher if the gas is denser. SNe deposit their energy in the ISM,

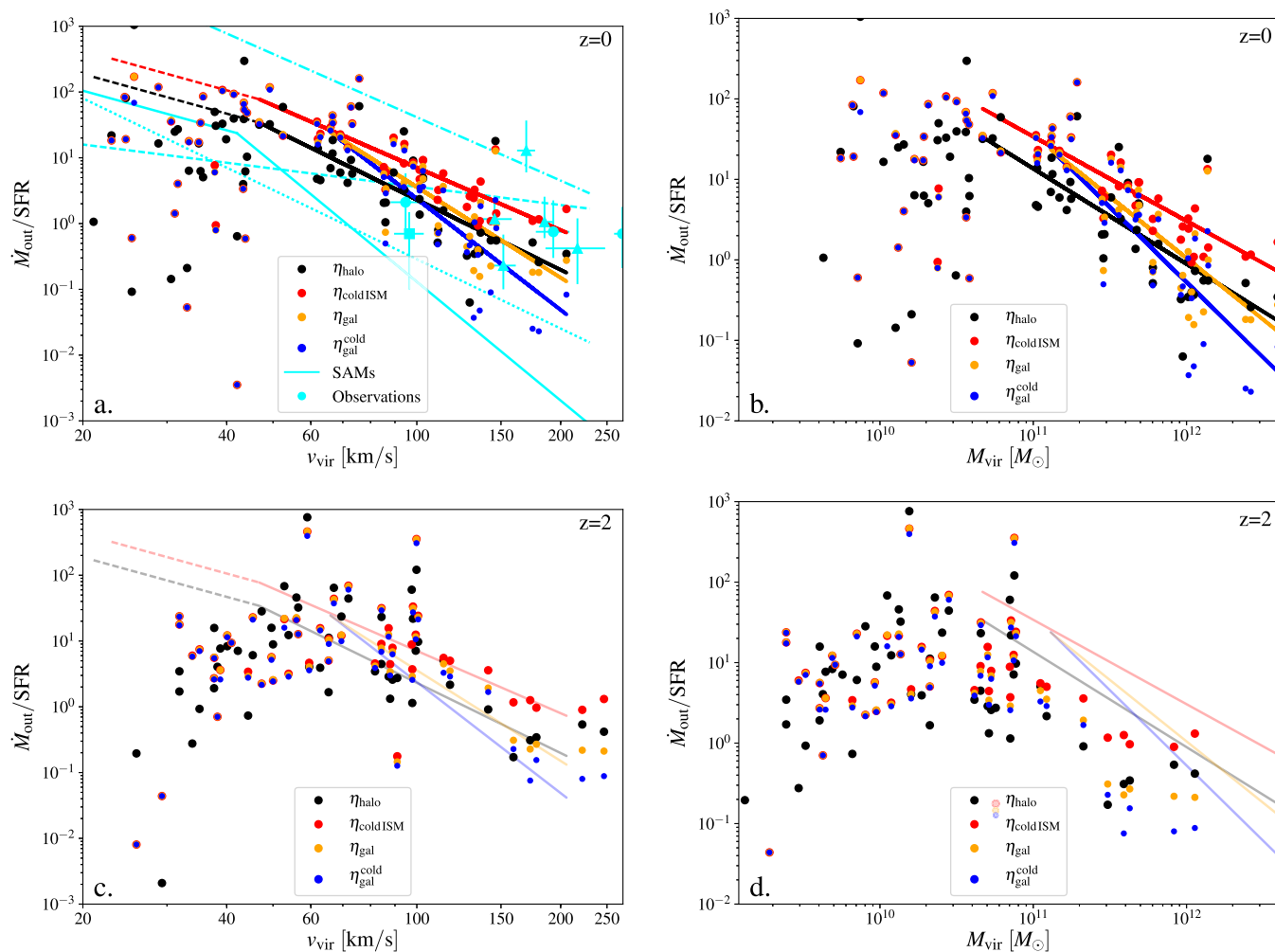


Figure 12. The mass-loading factors η_{coldISM} (red symbols), η_{gal} (yellow symbols), $\eta_{\text{gal}}^{\text{cold}}$ (blue symbols), and η_{halo} (black symbols). Mass-loading factors are shown as a function of both virial velocity v_{vir} (left) and virial mass M_{vir} (right). The upper and lower panels show the results of the NIHAO simulations at $z = 0$ and 2 , respectively. The thick red, blue, yellow, and black solid lines are log–log linear least-squares fits to the red, blue, yellow, and black symbols at $z = 0$, respectively. The red dashed line and the black dashed line correspond to relations of the form $\eta \propto v_{\text{vir}}^{-2}$ (as expected if feedback had a constant efficiency, e.g. 100 per cent). They have been repeated with a smaller linewidth on the panels at $z = 2$ to show the evolution of $\eta(M_{\text{vir}})$ and $\eta(v_{\text{vir}})$ with redshift. The cyan lines in the top left panel show the $\eta_{\text{gal}}-v_{\text{vir}}$ relations in four SAMs: Guo et al.’s (2011) version of the Munich model (dotted line), Henriques et al.’s (2013) version of the Munich model (dashed line), the Durham model (Lacey et al. 2016; dotted-dashed line) and GALICS 2.0 (Cattaneo et al. 2017; solid line). The cyan points with error bars in Fig. 12(a) are the observations from Bouché et al. (2012, triangles), Kacprzak et al. (2014, square), and Schroetter et al. (2015, circles). Schroetter et al. (2015) observe galaxies at $z \sim 0.8$, Kacprzak et al. (2014) at $z \sim 0.2$, and Bouché et al. (2012) at $z \sim 0.1$.

where even the dilute hot component is relatively dense. Only by expanding faster than it can cool can gas retain enough energy to produce an outflow. The hot CGM of massive galaxies confine the outflowing gas and slow down its expansion. Hence, the outflowing gas radiates a much larger fraction of its internal energy.

In Section 7, we shall present another physical reason why ϵ_{SN} should decrease with v_{vir} . However, the NIHAO simulations themselves show that this must be the case, since all fits in Table 1 decrease much faster than v_{vir}^{-2} . So do SAMs of galaxy formation, since nearly all of them require $\alpha \ll -2$ to be consistent with the slope of the stellar mass function of galaxies at low masses (see e.g. Guo et al. 2011; Lacey et al. 2016; Cattaneo et al. 2017, but also Henriques et al. 2013; cyan lines in Fig. 12a⁶).

Still, ϵ_{SN} cannot keep growing indefinitely at low v_{vir} because it cannot become larger than unity. Hence, there must be a virial velocity below which $\eta \leq \eta_{\text{max}} \propto v_{\text{vir}}^{-2}$ for all galaxies. Cattaneo et al. (2017) predicted this behaviour based on this consideration in the context of the GALICS 2.0 SAM and argued that it could explain the upturn of the stellar mass function of galaxies at low masses, which is observed both in the local Universe (e.g. Yang, Mo & van den Bosch 2009; Baldry et al. 2012; Moustakas et al. 2013) and at higher redshifts (e.g. Tomczak et al. 2014).

The NIHAO simulations confirm this prediction. All galaxies with $v_{\text{vir}} < 45 \text{ km s}^{-1}$ lie below the red dashed line $\eta_{\text{gal}} \simeq 90(v_{\text{vir}}/45 \text{ km s}^{-1})^{-2}$ in Fig. 12(a). This line corresponds to $\epsilon_{\text{SN}} = 1$ for $E_{\text{SN}} \Psi_{\text{SN}} = 8 \times 10^{48} \text{ erg M}^{-1}$ (Section 2) and $v_{\text{esc}}(r_g) \simeq 2.1 v_{\text{vir}}$,

⁶Fig. 12(a) shows that the normalization of η varies wildly from one SAM to another. GALICS 2.0 (Cattaneo et al. 2017) reproduces the stellar mass

function of galaxies with a lower normalization than other models because in GALICS 2.0, the ejected gas cannot be reaccreted.

which is a reasonable estimate for the escape speed from galaxies. Adding early stellar feedback increases the energy injected into the ISM by a factor of four. However, the fact that we find $\epsilon_{\text{SN}} = 1$ for a reasonable escape speed when we consider the contribution of SN feedback alone confirms the supposition that energy from early stellar feedback is dissipated much more efficiently and that its main role is to dilute the ISM in which SNe are about to explode.

The escape speed from the haloes is higher than the one from the galaxy. Hence, the outflow rate of the halo should be lower than the outflow rate from the galaxy by a factor $v_{\text{esc}}^2(r_{\text{g}})/v_{\text{esc}}^2(r_{\text{vir}})$. Assuming an NFW profile (Navarro, Frenk & White 1997), the escape speed to $r > r_{\text{vir}}$ from the centre of the halo is:

$$v_{\text{esc}}(r_{\text{vir}}) = \sqrt{\frac{2[c - \ln(1+c)]}{\ln(1+c) - c/(1+c)}} v_{\text{vir}}, \quad (9)$$

where c is the concentration of the NFW profile. The fitting formulae by Dutton & Macciò (2014) give $c \simeq 14$ for a halo with $v_{\text{vir}} = 45 \text{ km s}^{-1}$ in a cosmology with $\Delta_c = 109$ at $z = 0$ (Section 3.1). For this concentration, equation (9) gives $v_{\text{esc}}(r_{\text{vir}}) \simeq 3.6 v_{\text{vir}}$. Hence, we expect $\eta_{\text{halo}}/\eta_{\text{gal}} \sim (2.1/3.6)^2 \simeq 0.3$ at low virial velocities. Fig. 12(a) shows that this very crude estimate is not far off from reproducing correctly the ratio of the black dashed line to the red dashed line.

Fig. 12(a) shows the relation between η and v_{vir} . As $M_{\text{vir}} \propto v_{\text{vir}}^3$, it is not surprising that we find a similar relation between η_{galaxy} and M_{vir} (Fig. 12b). However, the proportionality constant between M_{vir} and v_{vir}^3 depends on redshift. Hence, the dependences of η on v_{vir} and M_{vir} cannot both remain the same at $z = 0$ and 2. At least one has to evolve. The one that evolves less is the more fundamental one. Figs 12(c) and (d) show that the η - v_{vir} relation is more or less the same at $z = 2$ and 0, whereas, for a same M_{vir} , η is lower at $z = 2$, since the escape speed for a same virial mass is higher at high redshift. This finding confirms the theoretical expectation that the gas's ability to escape is linked to the depth of the gravitational potential well more than it is to halo mass.

We remark that the picture presented in this section is in stark contrast with the one by Muratov et al. (2015, FIRE), who claimed that the logarithmic slope of $\eta(v_{\text{vir}})$ is steeper for ultradwarves ($\alpha = -3.2$), shallower for dwarves ($\alpha = -1$), and then very steep again for massive spirals. Muratov et al. (2015) attribute the change of exponent from $\alpha = -3.2$ to -1 around $v_{\text{vir}} = 60 \text{ km s}^{-1}$ to a transition from energy- to moment-driven winds. Had the NIHAO simulations included momentum feedback (e.g. radiation pressure) and not only energy feedback (thermal feedback from SNe and earlier phases of stellar evolution), we might have discovered an intermediate range of v_{vir} where $\eta(v_{\text{vir}})$ is shallower than what we find, although it is equally possible that scatter might have blurred any such effect (we have already remarked that Fig. 12b is consistent with a later FIRE publication by Anglés-Alcázar et al. 2017). We consider, however, that the exponent $\alpha = -3.2$ at $v_{\text{vir}} < 60 \text{ km s}^{-1}$ is unphysical, and that it derives from defining as outflowing any particle with positive radial velocity, even if its speed is lower than the escape speed from the galaxy, whereas the behaviour that we find is much more consistent with the episodic nature of star formation and outflows in ultradwarves.

We conclude this section by comparing the lessons from mass-loading factors and ejected fractions. In Section 4, we saw that $f_{\text{ej}}^{\text{gal}}$ decreases with M_{vir} faster than $f_{\text{ej}}^{\text{halo}}$ (Fig. 6). Hence, we may have expected that $\eta_{\text{cold ISM}}$ would decrease with v_{vir} faster than η_{halo} ($\eta_{\text{cold ISM}}$ is the relevant quantity for this comparison because the calculation of $f_{\text{ej}}^{\text{gal}}$ is based on M_{gal} , the total mass of stars and cold neutral gas within a galaxy). On the contrary, the α parameters for

$\eta_{\text{cold ISM}}$ and η_{halo} are very similar (Table 1) and $\eta_{\text{halo}} \sim \eta_{\text{galaxy}}/3$ on average. If anything, it is η_{halo} that decreases slightly faster (compare the red and the black lines in Fig. 12a). This apparent paradox arises because $f_{\text{ej}}^{\text{gal}}$ is linked to the mass $M_{\text{gal,acc}} - M_{\text{gal}}$ of the baryons that are no longer in the galaxy at $z = 0$ but have been inside it at some earlier epoch, while $\eta_{\text{cold ISM}}$ is linked to the instantaneous rate $\dot{M}_{\text{out, cold ISM}}$ at which gas leaves the cold ISM. If gas that leaves the cold ISM never came back, then:

$$M_{\text{gal,acc}} - M_{\text{gal}} = \int \dot{M}_{\text{out, cold ISM}} dt, \quad (10)$$

but we shall see that this is not the case (Section 8). Hence, the different behaviour of $f_{\text{ej}}^{\text{gal}}$ and $\eta_{\text{cold ISM}}$ must be due to reaccretion: $f_{\text{ej}}^{\text{gal}}$ decreases more rapidly than $\eta_{\text{cold ISM}}$ because, in massive galaxies, a lot of gas is ejected and reaccreted. This gas contributes to $\eta_{\text{cold ISM}}$ but not to $f_{\text{ej}}^{\text{gal}}$.

7 OUTFLOW SPEED AND MULTIPHASE STRUCTURE

In Section 6, we investigated the rate at which gas is expelled from galaxies and we saw that separating the cold component from the total outflow rate is important for an accurate comparison with the observations (Fig. 11). Here, we analyse the outflow speed of each component by studying its distribution of radial velocity v_r , so that we can separate inflows ($v_r < -v_{\text{vir}}$), outflows ($v_r > v_{\text{vir}}$), and hot gas in hydrostatic equilibrium or cold galactic gas ($|v_r| \ll v_{\text{vir}}$).

Fig. 13 shows the distribution of radial velocity v_r for the cold gas (cold ISM plus cold IGM), the hot ISM, and the hot CGM of the four galaxies listed in Table 2. Galaxies g1.23e10 and g1.90e10 typify the bimodality of star formation in ultradwarves, which can be very inefficient (as in g1.23e10) or very intense (as in g1.90e10). Galaxies g3.21e11 and g1.12e12 are representative of massive dwarves/low-mass and massive spirals, respectively, which constitute a more homogeneous population than ultradwarves, even though significant differences from one galaxy to another are observed.

The first result that emerges from Fig. 13 is that the presence of a hot ISM is linked to recent star formation. The passive ultradwarf g1.23e10 displays neither outflows nor a hot ISM. In contrast, a hot ISM is present in the other three galaxies, all of which display star formation and outflows.

Secondly, when dwarf galaxies do exhibit outflows (as in g1.90e10), they are normally dominated by the cold component. At higher masses, the hot dense component carries most of the gas with $v_r > 200 \text{ km s}^{-1}$. This is the same result that we had already found under a different guise in Figs 11 and 12. Notice, however, that the bulk of the hot dense gas is not outflowing (all galaxies in Table 2 have $\bar{v}_r^{\text{hot ISM}} < 0.2 v_{\text{esc}}$). This is why we call this phase hot ISM and not wind.

Finally, the distributions of outflow speeds in g3.21e11 and g1.12e12 are not very different, but the escape speeds are. The vertical dashed lines in Fig. 13 show the escape speeds from r_{vir} computed with equation (9), but the argument that follows would still apply if we considered escape speeds from r_{g} instead. In g3.21e11, the gas fraction with $v_r > v_{\text{esc}}$ is small (relative to the total mass of gas)⁷ but still significant. In g1.12e12, no outflow is

⁷The observation that the mass in the wind is always small compared to the mass of the galaxy does not preclude the possibility that most of the baryons may be ejected. There are many more healthy people than sick ones. Yet, most of us will fall ill and die at some point.

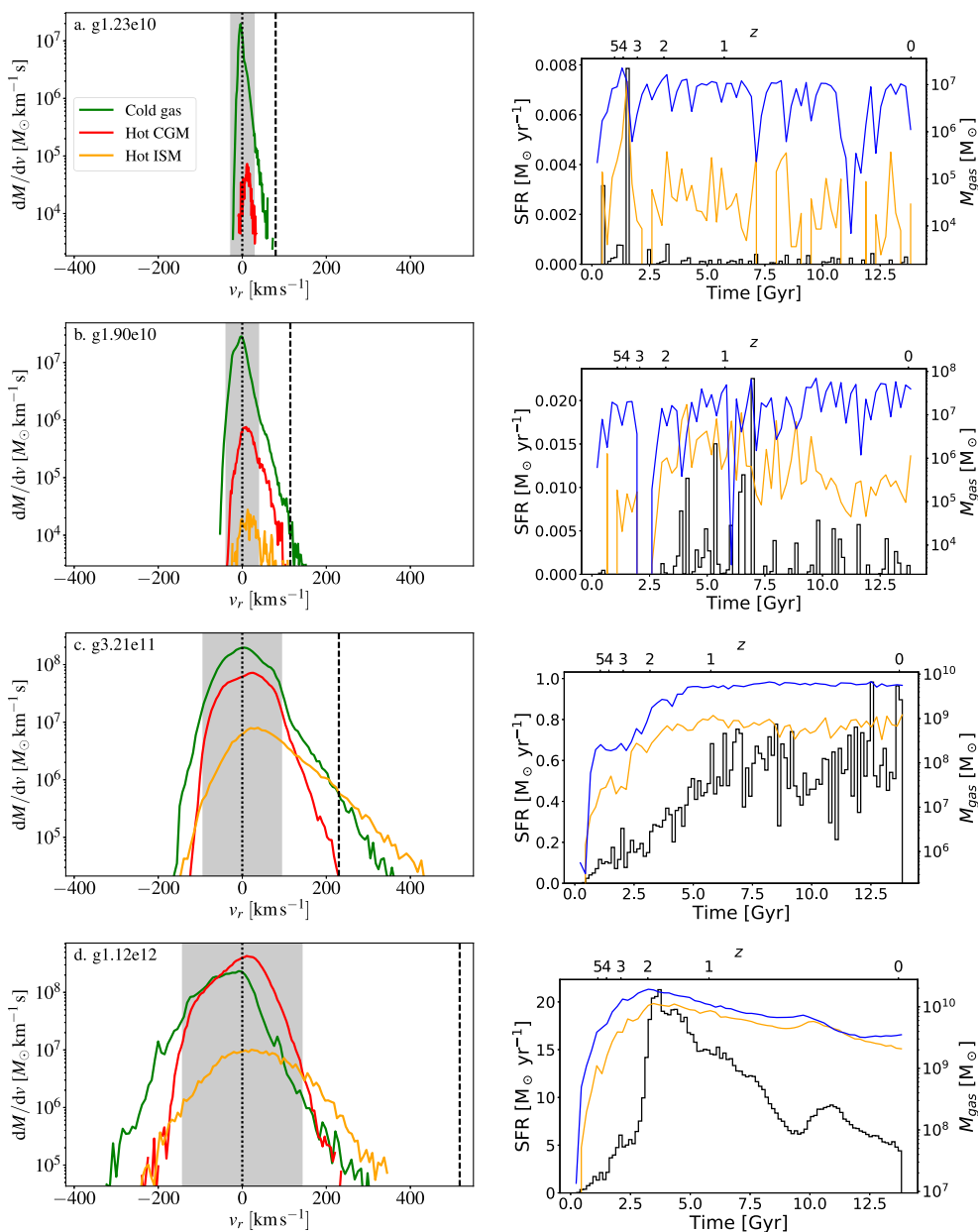


Figure 13. Left: radial-velocity distribution of the cold gas (ISM and CGM, green curves), the hot ISM (yellow curves) and the hot CGM (red curves) for the four objects in Table 2 arranged by growing M_{vir} . The normalizations of the curves reflect the masses in the three phases (the ordinate range is not the same in all panels). The vertical dashed lines show the escape speeds v_{esc} from the centres of the haloes. The shaded areas corresponds $-v_{\text{vir}} < v_r < v_{\text{vir}}$. Right: corresponding SFR histories (black histograms) and evolution of the mass in the hot ISM (yellow curves) and the cold CGM (blue curves). If the SFR rises, so does the mass in the hot ISM. Notice that we use a linear scale SFRs and a logarithmic one for $M_{\text{hot ISM}}$.

Table 2. Virial masses, stellar masses, escape speeds, and mean radial velocities of the cold gas, the hot CGM and the hot ISM for the four NIHAO galaxies in Fig. 13. All speeds are in km s^{-1} .

Name	$M_{\text{vir}} [M_{\odot}]$	$M_* [M_{\odot}]$	v_{esc}	\bar{v}_r^{cold}	$\bar{v}_r^{\text{hot ISM}}$	$\bar{v}_r^{\text{hot CGM}}$
g1.23e10	1.05×10^{10}	1.60×10^6	79.4	-0.85	Not detected	12.0
g1.90e10	2.70×10^{10}	1.20×10^7	114	-2.61	23.1	15.2
g3.21e11	3.61×10^{11}	3.73×10^9	230	6.67	57.0	13.6
g1.12e12	1.29×10^{12}	7.95×10^{10}	517	-41.0	32.4	-3.66

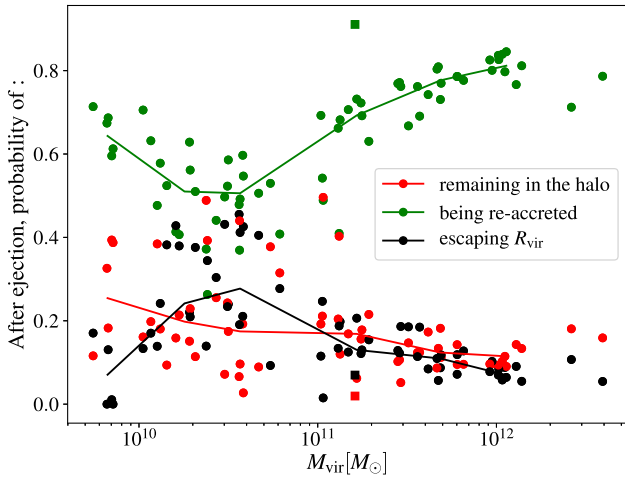


Figure 14. Fraction of the ejected mass from the galaxy that escapes from the halo and never comes back (black symbols), remains in the halo (red symbols) and accretes back on to the galaxy (green symbols). Symbols refer to individual NIHAO galaxies. Their sum is unity. The curves are median values in bins of M_{vir} . Each curve corresponds to the symbols of the same colour.

detected at $z = 0$. In massive galaxies, SNe waste a lot of energy accelerating particles that will not escape. Since the efficiency of SNe that is relevant for the calculation in equation (8) refers to the kinetic energy of particles with $v_r > v_{\text{esc}}$, this is another reason why ϵ_{SN} decreases with v_{vir} in addition to the appearance of confining atmospheres at high masses discussed in Section 6 (comparing Figs 13c and d shows that the hot CGM is more important in g1.12e12 than it is in g3.21e11).

8 THE BARYON CYCLE

Having discussed accretion (Section 4), star formation (Section 5), and outflows (Sections 6 and 7), we are now ready to address the fate of the ejected gas and close the loop of the galaxy baryon cycle.

Fig. 14 anticipates our most basic result. In dwarves with $M_{\text{vir}} \sim 2 - 4 \times 10^{11} M_{\odot}$, out of $100 M_{\odot}$ expelled from the galaxy, ~ 30 per cent escapes from the halo and never comes back, ~ 20 per cent remains inside the halo until $z = 0$, and ~ 50 per cent accretes back on to the galaxy, although some of them will be blown out again. In massive galaxies with $M_{\text{vir}} > 3 \times 10^{11} M_{\odot}$, the reaccreted fraction rises to ~ 80 per cent and less than 10 per cent of the gas is expelled from the halo. Thus, reaccretion is statistically the single most likely outcome at all halo masses but especially at the highest. Our recycled fractions (green curve in Fig. 14) are consistent with the recycled fractions at $z = 0$ from fig. 9 of Anglés-Alcázar et al. (2017).

We have stressed the possibility that some of the reaccreted gas may be ejected again because all figures in this paragraph are for a single ejection event. If a particle is ejected several times, the probability that it is expelled from the halo once will be larger than the 10 per cent–30 per cent figure above.

To go beyond this final outcome, and study the pathways through which gas is expelled and reaccreted, we begin with our favourite galaxy g1.12e12 and consider all particles within the virial radius at the penultimate output time $z = 0.05$ (Fig. 15), starting from those in the hot CGM.

The hot CGM that cools in the time interval $\Delta t = 0.216$ Gyr between $z = 0.05$ and 0 is about 1 per cent of the mass in the hot CGM (Fig. 15a). From this, one infers a cooling time of more than 20 Gyr, much longer than the one expected from the density, temperature, and metallicity of the hot CGM (the cooling function by Sutherland & Dopita 1993 gives ~ 0.7 Gyr). The difference arises because the cooling time that we infer from Fig. 15(a) is not $(3/2)nkT/C$, where C is the power radiated per unit volume, but rather $(3/2)nkT/C - \mathcal{H}$, where \mathcal{H} is the thermal power that SNe transfer to hot CGM per unit volume.

To test this explanation, we have computed the total power L_X radiated by CGM and we have compared it to the power $P_{\text{SN}} = E_{\text{SN}}\Psi_{\text{SN}}\text{SFR}$ from SNe, where $E_{\text{SN}} = 10^{51}$ erg, and $\Psi_{\text{SN}} = 1/(120 M_{\odot})$ (we consider the power from SNe only and neglect the contribution from pre-SN feedback because we assume that the latter is dissipated already at the level of the ISM). For galaxy g1.12e12 at $z = 0$, we find $L_X = 1.2 \times 10^{42}$ erg s^{-1} and $P_{\text{SN}} = 1.6 \times 10^{42}$ erg s^{-1} . Both quantities are likely to be overestimated: the NIHAO simulations use the cooling function by Shen et al. (2010), which considers the presence of a UV background and gives lower cooling rates than the one by Sutherland & Dopita (1993); the thermal power that SNe inject into the CGM will be lower than P_{SN} because some of the energy from SNe, too, may be dissipated at the level of the ISM and not reach the CGM. Furthermore, P_{SN}/L_X varies stochastically from one snapshot to the next. However, within these uncertainties, our calculations support the conclusion that in massive spirals, such as g1.12e12, the cooling and the heating rates comparable, and thus balance each other off on average. This explains the very low cooling rate that we infer from Fig. 15(a). We remark that this maintenance feedback is equivalent to the one by radio galaxies in clusters (e.g. Birzan et al. 2004; Rafferty et al. 2006; Cattaneo et al. 2009, and references therein), only spiral galaxies have more star formation and lower mass black holes than elliptical galaxies, hence it is SNe that heat the gas.

Physically, SNe heat the hot CGM via two mechanisms: the hot ISM vents out of the galaxy, mixes with the hot CGM, and transfers its thermal energy to it; SN-driven outflows from the galaxy shock heat the hot CGM. The cold CGM is located mainly in the plane of the disc and galactic wind mixes with the cold CGM only after radiating their thermal energy (in the descending part of the fountain). Hence, SN heating couples much more effectively to the hot CGM phase than to the cold CGM phase.

The prevalent motion in the cold CGM is accretion on to the galaxy (Fig. 15b). Particles that leave the cold ISM are, in their vast majority, particles that have been heated by SNe and transferred to the hot ISM (Fig. 15c). The hot ISM is a highly unstable phase that needs constant replenishment. By $z = 0$, more than half of the particles in the hot ISM at $z = 0.05$ have either cooled (43 per cent) or they have moved to the hot CGM (13 per cent; Fig. 15d). Hence, cooling is thrice more probable than becoming part of the hot CGM (twice more probable if we consider only cooling *in situ*, that is, within the ISM).

These results are based on instantaneous mass exchange rates for one galaxy in our sample, which is not a random object but the second most massive galaxy of the NIHAO series. To extend and generalize these findings to the entire NIHAO series, we have computed the mean lifetimes in the hot ISM, the cold CGM, the hot CGM, and the IGM ($r > r_{\text{vir}}$) for particles that entered these phases after being ejected from the cold ISM (Fig. 16). By construction, these lifetimes cannot be longer than the age of the Universe and they

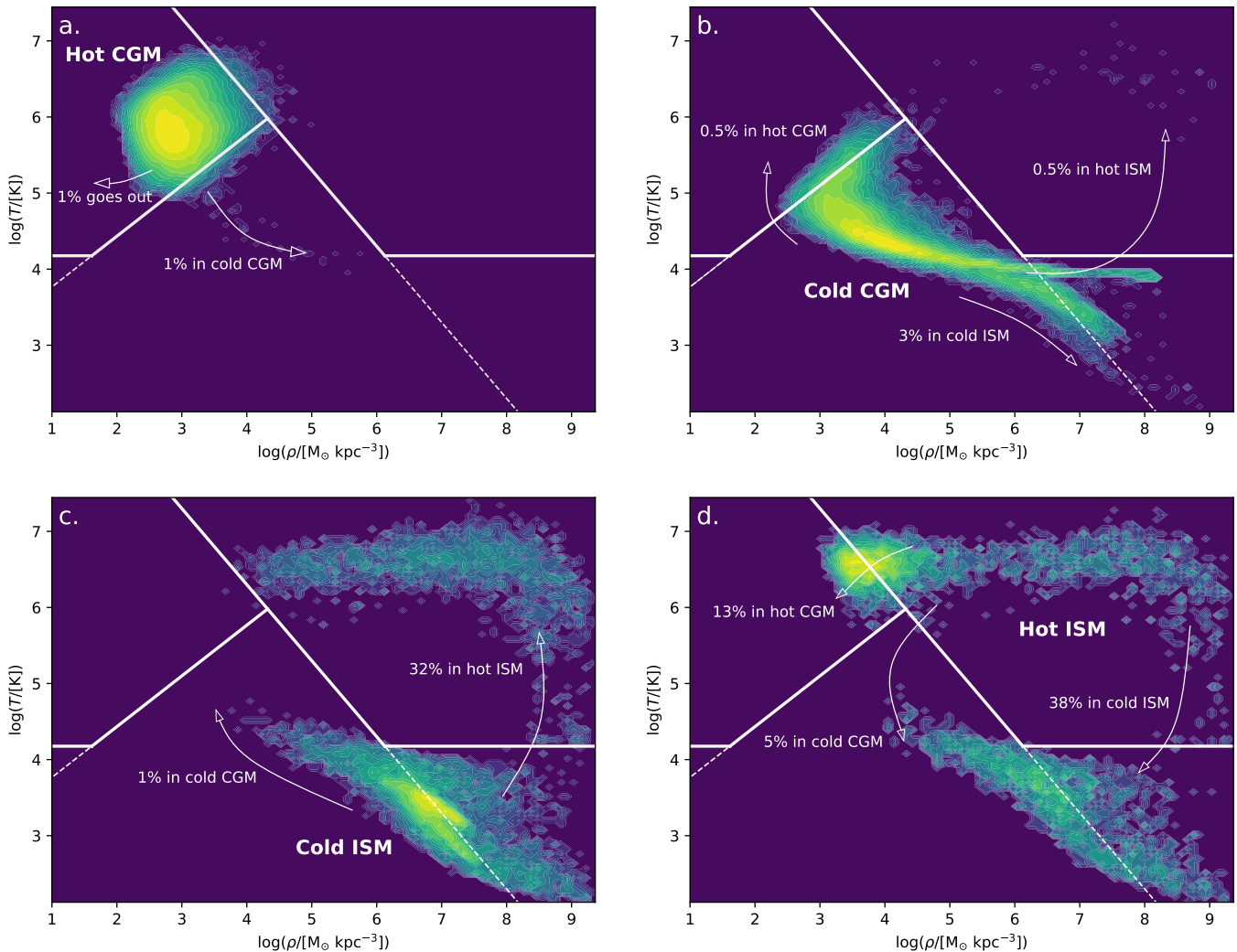


Figure 15. Panels (a)–(d) show the position on a temperature–density diagram at $z = 0$ of particles that were in the hot CGM, the cold CGM, the cold ISM, and the hot ISM at $z = 0.05$, respectively. This figure is for the NIHAO simulation g1.12e12. Percentages are relative to the mass of the phase from which a transition starts and to the time interval between $z = 0.05$ and 0 (0.216 Gyr).

cannot be shorter than the time interval $\Delta t = 0.216$ Gyr between two consecutive outputs (dashed horizontal line).

Fig. 16 shows that the hot ISM is so short-lived that its lifetime is not adequately resolved by our analysis. To quantify the impact that this may have on our results, we have also considered one galaxy (g1.37e11) that was simulated with much closer output time-steps. Using $\Delta t = 13$ Myr allows us to resolve the time spent in the hot ISM, which is larger than 30 Myr by construction, since we model SN feedback as in Stinson et al. (2006).

In Fig. 16, the mean lifetimes for g1.37e11 are shown with open squares rather than open circles. The lifetime in the hot ISM is so short (37 Myr) that the yellow square is outside the plotting range: as soon as gas is allowed to cool, it does it in just 7 Myr. The blue open square and the red open square lie a factor of two below the blue and the red curves, respectively, but their orders of magnitude are reproduced correctly. The black open square is consistent with the black curve.

Our findings for the blue open square and the red open square should not surprise us. Consider a particle that it is in a same

phase at the consecutive output times t_1 and $t_2 > t_1$; we conclude that this particle has been in that phase for a time $\geq t_2 - t_1$. If a snapshot at an intermediate time t_3 becomes available, the additional information at t_3 may support this conclusion but also disprove it (if we find that at t_3 the particle is in a different phase). Hence, a finer time resolution can only shorten the mean time spent uninterruptedly in any phase, and this is what Fig. 16 shows. However, the integrated total time spent in any phase robust to this effect.

With the caveat above in mind, Fig. 16 shows that gas expelled from haloes remains at $r > r_{\text{vir}}$ for a very long time (typically, until the simulation ends). In contrast, cold gas within haloes is usually reaccreted on a time-scale of order ~ 1 Gyr, which is shorter at high M_{vir} . This behaviour has a simple physical explanation. Cold outflows are approximately ballistic. Hence, the reaccretion time-scale t_{reaccret} is simply the dynamical time $r_{\text{vir}}/v_{\text{vir}}$ (in our cosmology, $r_{\text{vir}}/v_{\text{vir}} = 1.9$ Gyr at $z = 0$) times a factor that depends on v_0/v_{vir} and concentration only, if we assume that the DM dominates the gravitational potential (v_0 is the initial speed at which gas

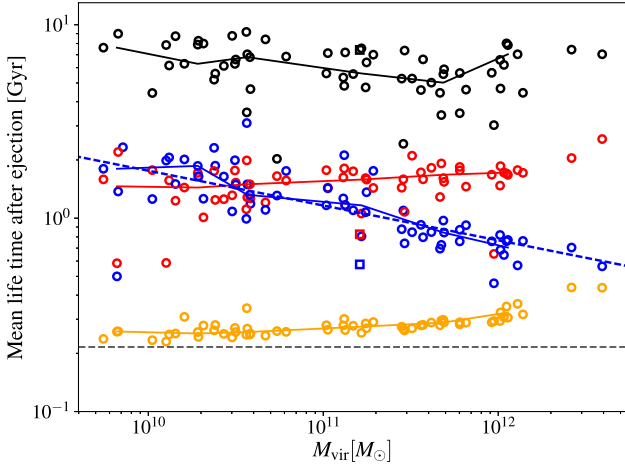


Figure 16. Mean mass-weighted lifetime in the IGM (gas outside the virial radius, black symbols), the hot CGM (red symbols), the cold CGM (blue symbols), and the hot ISM (yellow symbols) for particles that enter these phases after leaving the cold ISM (particles are not required to reaccrete on to the galaxy to be counted for this figure). The solid lines show the median behaviour of the symbols of matching colour. The dotted black line shows the time resolution of our analysis (0.216 Gyr at all z). The open squares show the same results for a simulation that was analysed with a time resolution of 0.013 Gyr. In this simulation, the lifetime in the hot ISM is 37 Myr. The blue dotted line is a fit of the blue circles.

is ejected).⁸ Differences due to concentration are small and the dynamical time is the same for all haloes at a given redshift. Thus, if v_0/v_{vir} were the same for all haloes, t_{reacrer} would be the same, too. If, on the contrary, v_0 were constant, then t_{reacrer} would be a strongly decreasing function of v_{vir} . Its form would be more complicated than a simple power law of v_{vir} . However, if we assume an NFW gravitational potential with $c \sim 11\text{--}14$, a universal outflow speed implies $t_{\text{reacrer}} \propto v_{\text{vir}}^{-3}$ for $0.4 < v_{\text{vir}}/v_0 < 0.8$ and $t_{\text{reacrer}} \propto v_{\text{vir}}^{-2}$ for $v_{\text{vir}}/v_0 \gtrsim 0.8$ ($v_{\text{vir}}/v_0 = 0.8$ corresponds to $t_{\text{reacrer}} \sim 0.4$ Gyr at $z=0$). Fig. 13 shows an intermediate situation: v_0 increases with v_{vir} but not as rapidly as v_{vir} . Hence, the reaccretion time decreases with M_{vir} but not as rapidly as assumed by Henriques et al. (2013, 2015), who used $t_{\text{reacrer}} \propto M_{\text{vir}}^{-1}$ independent of redshift.

If we fit the blue circles in Fig. 16 with a power-law function of M_{vir} , we find:

$$t_{\text{reacrer}} \simeq 1.1 \left(\frac{M_{\text{vir}}}{10^{11} M_{\odot}} \right)^{-0.18} \text{ Gyr.} \quad (11)$$

(blue dashed line in Fig. 16), although it is possible that equation (11) may overestimate the reaccretion time for cold outflows by up to a factor of two (see above for a discussion of how time resolution can affect our results). Equation (11) gives $t_{\text{reacrer}} \sim 0.7$ Gyr for $M_{\text{vir}} = 10^{12} M_{\odot}$, which becomes $t_{\text{reacrer}} \sim 0.4$ Gyr if

⁸In an NFW halo:

$$t_{\text{reacrer}} = 2 \frac{r_{\text{vir}}}{v_{\text{vir}}} \int_0^{x_{\text{max}}} \frac{dx}{\sqrt{2 \frac{c}{f(c)} \frac{\ln(1+x)}{x}}},$$

where x_{max} is found by solving the energy-conservation equation:

$$\frac{1}{2} \left(\frac{v_0}{v_{\text{vir}}} \right)^2 = \frac{c}{f(c)} \left[1 - \frac{\ln(1+x_{\text{max}})}{x_{\text{max}}} \right]$$

and $f(c) = \ln(1+c) - c/(1+c)$.

we renormalize the blue dashed line so that it passes through the blue square. These reaccretion times are intermediate between those in Oppenheimer et al. (2010) and Christensen et al. (2016; $t_{\text{reacrer}} \sim 1$ Gyr), and those in Anglés-Alcázar et al. (2017; $t_{\text{reacrer}} \sim 0.1 - 0.35$ Gyr).⁹ For comparison, the semiempirical model by Mitra, Davé & Finlator (2015) reproduces the observations with $t_{\text{reacrer}} \propto 0.52 (M_{\text{vir}}/10^{12} M_{\odot})^{-0.45} (1+z)^{-0.32}$ Gyr.

Discussing the lifetime in the hot CGM requires more subtlety. Hot outflows can be reaccreted only if they cool down and it is well known (since Rees & Ostriker 1977; Silk 1977; Binney 1977) that the cooling time is longer at higher masses. However, if we interpret the red curve in Fig. 16 as a measure of the radiative cooling time, then Fig. 16 gives results that conflict with our estimate for g1.12e12 (Fig. 15a). The reason is that gas can leave the hot CGM not only because it cools but also and more likely because it escapes from the halo (hot outflows). Hence, the typical 1 – 2 Gyr lifetime depicted by the red curve in Fig. 16 measures the time to leave the virial radius (a particle with an average speed of 100 km s^{-1} travels 100 kpc in about 1 Gyr) more than the cooling time.

This conclusion may appear to conflict with our previous finding that most of the hot CGM has radial velocities lower than the escape speed (Fig. 13). However, the escape speed in Fig. 13 is the escape speed from the centre of the halo, while most of the hot CGM is at several galactic radii.

Massive spirals such as g1.12e12 are the only ones that have proper CGM composed of hydrostatic rather than outflowing gas and are not representative of the average NIHAO galaxy in this regard. For these massive spirals, we have recomputed the mean lifetime in the hot CGM when we retain only particles that cool and return to the ISM.¹⁰ We expected that this calculation would give us a longer time than the one we found by considering all particles. On the contrary, when we applied this analysis to g1.12e12 the survival time in the hot CGM decreased from 2 to 0.7 Gyr (in even greater disagreement with the cooling time that we infer from Fig. 15a). Our interpretation is that Fig. 15(a) gives the correct cooling time for the hot CGM of g1.12e12, but this cooling time is an average over two types a particles: a few particles that cool within the radiative cooling time $(3/2)nKT/C \sim 0.7$ Gyr expected if there were no heating and many more that never do.

The final step of our analysis is to study the different paths through which baryons can be ejected and reaccreted and to compute the masses processed through each of them. For this purpose, we introduce the concept of cycle. A cycle begins when a particle leaves the cold ISM. It finishes when the particle returns to the cold ISM (closed cycle) or when the simulation reaches $z = 0$ (open cycle), whichever occurs first.

The frequency of a type of cycle (e.g. the cycle cold ISM \rightarrow hot ISM \rightarrow hot CGM \rightarrow cold CGM \rightarrow cold ISM) is quantified by the mass processed through the cycle divided the total mass $M_{\text{gal,acc}}$ processed by the galaxy throughout the lifetime of the Universe. If all particles that accrete on to the galaxy went through this cycle once, then the cycle's frequency would be unity. This is impossible because some particles must form stars for feedback to occur. However, a cycle's frequency could still be equal to one

⁹Anglés-Alcázar et al. (2017) used a time resolution of 0.1 Gyr, higher than the one used to compute the blue circles in Fig. 16 (0.216 Gyr) but lower than the one used to compute the open square (0.013 Gyr).

¹⁰We stress that this lifetime gives a cooling time and not a reaccretion time because after cooling gas must still fall back on to the galaxy. Hence, the reaccretion time will be longer.

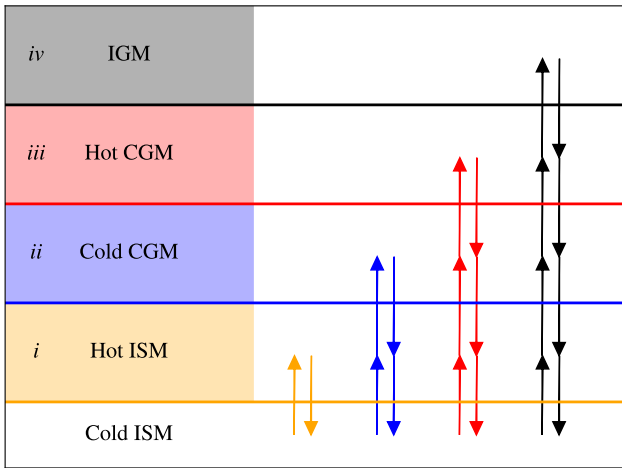


Figure 17. Cycles considered in Fig. 18. All cycles start in the cold ISM. They must pass through the cold ISM, the cold CGM, the hot CGM and the IGM to be classified as type *i*, *ii*, *iii*, or *iv*, respectively, but may miss intermediate phases. Cycles are subdivided into closed and open. Closed cycles return to the cold ISM. Open cycles may or may not have a descending branch but have not returned to the cold ISM by $z = 0$.

if half of the accreted particles formed stars and the other half went through the cycle twice. The cycle must be close for that to happen.

The hot ISM that vents out of galaxies may end up in either the hot or the cold CGM, and the hot CGM may become cold CGM, but gas cannot move from the cold CGM to the hot ISM or the hot CGM (at least not without being reaccreted on to the galaxy).¹¹ Neither is it physically possible that gas escapes from the galaxy to the ISM without passing through either the cold or the hot CGM, although it can happen in our analysis if the time resolution is inadequate. Hence, cycles can come in many different permutations, but there is a hierarchy of phases that allows to reduce them to four basic types (Fig. 17): (*i*) the cold ISM \rightarrow hot ISM \rightarrow cold ISM cycle, (*ii*) the cold ISM \rightarrow hot ISM \rightarrow cold CGM \rightarrow cold ISM cycle (and its variant without the hot ISM), (*iii*) the cold ISM \rightarrow hot ISM \rightarrow hot CGM \rightarrow cold CGM \rightarrow cold ISM cycle (and its variants without the hot ISM and/or the cold CGM), and (*iv*) any cycle that passes through the IGM. This classification has been formulated for closed cycles but we can easily generalize it to include open cycles by dropping the requirement that particles should return to the cold ISM.

Fig. 18 shows the average number of type *i*, *ii*, *iii*, and *iv* cycles that an average particle makes as a function of M_{vir} . The solid curves are for closed cycles only. The dashed curve include open cycles. The yellow solid and the yellow dashed curves run close to each other. So do the blue solid and the blue dashed curves. The implication is that types *i* and *ii* cycles are predominantly closed. They correspond to internal heating of the ISM by SNe and the galactic fountain, respectively. In contrast, the red solid and the red dashed curves are separated by about an order of magnitude. So are the black solid and the black dashed curves. Therefore, types *iii* and *iv* cycles are predominantly open (gas that mixes with the hot CGM or leaves the halo is unlikely to be reaccreted).

¹¹In reality, some particles do (Fig. 15a) but these transitions are not significant.

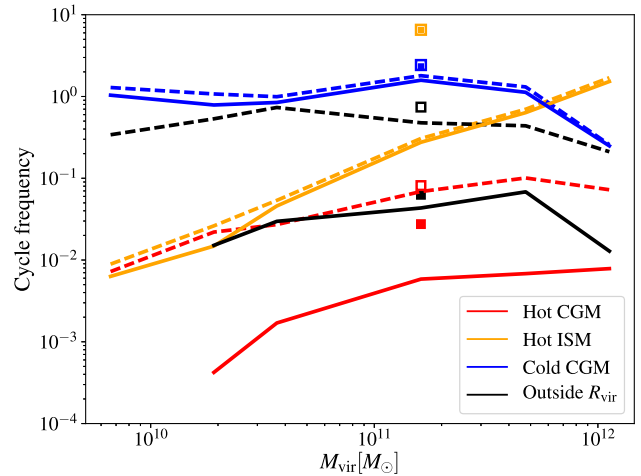


Figure 18. Mass-weighted frequency of the different cycles an accreted particle could follow (as per Fig. 17). All lines are medians based on following cycles with a time resolution of 0.216 Gyr. Solid lines are for closed cycles. Dashed lines include the contributions of both close and open cycles. Filled/open squares are equivalent to solid/dashed curves, respectively, for the galaxy in which the baryon cycle has been followed with a resolution of 0.013 Gyr.

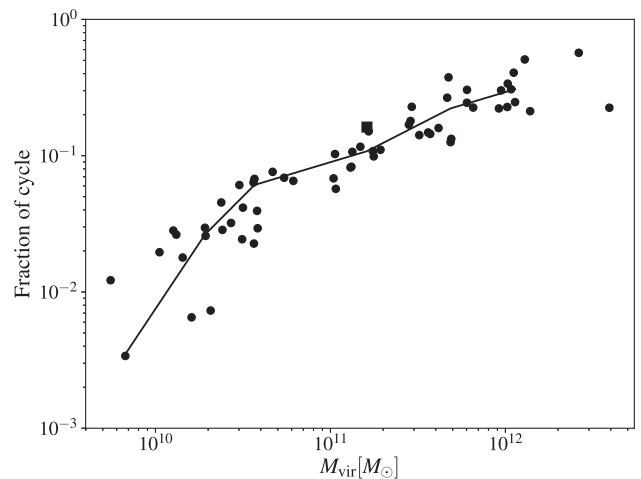


Figure 19. Fraction of type *ii* cycles with a transient hot-ISM phase. The square is for the simulation with snapshots separated by $\Delta t = 0.013$ Gyr.

Type *ii* cycles cold ISM \rightarrow (hot ISM) \rightarrow cold IGM \rightarrow cold ISM are the most common at all but the highest masses (Fig. 18, blue solid curve). An average gas particle makes one or two of them after entering the galaxy (Anglés-Alcázar et al. 2017, too, find that an average gas particle is ejected and reaccreted twice). The intermediate hot-ISM phase is missing in most cases but especially at low masses (Fig. 19). Hence, type *ii* cycles trace cold material entrained by the less massive hot component (also see Fig. 3) rather than hot gas that has expanded and cooled. As the galaxy analysed with $\Delta t = 0.013$ Gyr (black square in Fig. 19) follows the same mean relation as the other NIHAO galaxies, we are confident that Fig. 19 depicts a real physical result rather than our inability to detect short spells in the hot ISM.

Type *iv* cycles are the second most common variety. Their median frequency is about half per particle (Fig. 18, black dashed curve). This result does not conflict with Fig. 14, which shows that only 10 per cent–30 per cent of the ejected particles escape to $r > r_{\text{vir}}$,

because Fig. 14 refers to one ejection event, whereas Fig. 18 refers to the lifetime of the Universe. If we roll a dice once, the probability of an ace is one-sixth. If we roll it several times, the probability to obtain an ace at least once will be larger.

Type *i* cycles are rare at low masses but become predominant at $M_{\text{vir}} > 6 \times 10^{11} M_{\odot}$ (Fig. 18, yellow solid curve). Type *iii* cycles are rare at all masses but particularly at the lowest. Even in massive spirals, they affect at most one particle out of 10 (Fig. 18, red dashed curve).

As we have seen that the time resolution of our analysis is a potentially important issue, we have compared the curves in Fig. 18 with our findings for g1.37e11, for which we have a snapshot every 13 Myr. Better time resolution allows us to detect a much greater number of type *i* cycles (compare the yellow squares and the yellow curves), but these cycles are so short that, even collectively, their contribution is not significant anyway. If it were, the hot ISM would contain a significant fraction of the ISM's total mass. On the contrary, in the NIHAO simulations, the hot phase make up $\lesssim 10$ per cent of the mass of the ISM on average, although the value can rise to $\gtrsim 20$ per cent for massive spirals such as the Milky Way. The frequencies of the cycles that matter most (those of types *ii* and *iv*) appear to be robust. So are the frequencies of type *iii* cycles when we lump closed and open cycles together. The only difference is the frequency of type *iii* closed cycles. Type *iii* cycles are more likely to be closed in the simulation analysed with $\Delta t = 0.013$ Gyr than they are on average in the simulations analysed with $\Delta t = 0.216$ Gyr (the red filled square is much closer to the red open square than the red solid line is to the red dashed line). However, type *iii* closed cycles are so rare that this difference concerns 2 per cent of the baryons at most.

Fig. 18 is very important because it summarizes and quantifies the two main ways through which the expulsion of matter reduces the masses of galaxies: fountain and ejective feedback. In a typical galaxy, baryons are ejected and reaccreted once or twice and, in the process, 20 per cent–80 per cent (thus, half on average) are blown out of the virial radius and never come back.

Fountain feedback acts by reducing the time gas spends in the cold ISM (the only phase where it may form stars). To quantify this effect, we have considered all gas particles that accrete on to a galaxy from when they were first accreted to their last time in the cold ISM, which can be when they were last ejected, when they formed stars, or when the simulation ended. Fig. 20 splits this time interval into two: the time spent in the cold ISM (black circles) and the time spent in closed cycles outside the cold ISM (red circles). In dwarf galaxies with $10^{10} M_{\odot} < M_{\text{vir}} < 10^{11} M_{\odot}$, an average gas particle spends ~ 2.5 Gyr in the fountain and less than 1 Gyr in the cold ISM. Hence, it is available for star formation for only a quarter of the time. At $M_{\text{vir}} > 3 \times 10^{11} M_{\odot}$, the time available for star formation becomes $\gtrsim 50$ per cent but the time spent in the fountain remains significant. The galaxy that we have analysed at intervals of $\Delta t = 0.013$ Gyr spends ~ 10 per cent less time in the fountain than an average galaxy of the same mass (based on our analysis with $\Delta t = 0.216$) but the difference is so small that it will not change anything in our conclusions.

9 DISCUSSION AND CONCLUSION

The last decades have seen considerable progress in our understanding of the cosmological context in which galaxies form and evolve. These advancements are due to precise measurements of the cosmological parameters (Bennett et al. 2013; Planck Collaboration et al. 2014), galaxy surveys of unprecedented size and depth

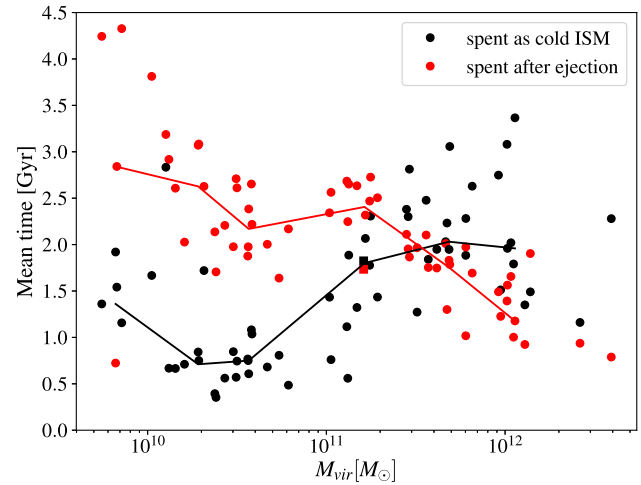


Figure 20. Mean time spent in the cold ISM (black circles) and the fountain (red circles) for gas particles that accrete on to galaxies. The low M_{*}/M_{vir} ratios of galaxies with $10^{10} M_{\odot} < M_{\text{vir}} < 10^{11} M_{\odot}$ are partly due to the short time spent in the cold ISM. The galactic fountain reduces the time available for star formation by a factor of 2–4.

(SDSS, 2dF, DEEP2, VVDS, UltraVISTA, and CANDELS), and hydrodynamic computer simulations (Hopkins et al. 2014; Schaye et al. 2015; Vogelsberger et al. 2014) but, even with today's most powerful supercomputers, the range of scales that a simulation can explore is still limited. Hence, feedback – the way processes on the scale of stars and the ISM affect the evolution of galaxies at large – has emerged as the main outstanding problem in galaxy formation.

All simulations need subgrid models to capture the physics they cannot resolve, but some models are more physical than others. The FIRE collaboration (Hopkins et al. 2014) has run simulations that include type I and type II SNe, radiation pressure, stellar winds, photoionization, and photoelectric heating. We cannot claim that our simulations follow the interaction of stars with the ISM to a comparable level, but neither is that our objective. Our goal is not to study the physics of stellar feedback but rather to explore their effects on the formation and evolution of galaxies.

The NIHAO simulations (Wang et al. 2015) are perfectly suited for that purpose because they contain 90 galaxies that span the entire mass range from $M_{\text{vir}} = 4 \times 10^9$ to $4 \times 10^{12} M_{\odot}$, and they have been shown to reproduce several properties of galaxies, which include stellar masses and SFRs (Wang et al. 2015), disc sizes and the Tully–Fisher relation (Dutton et al. 2017), galaxy velocity function (Macciò et al. 2016), and the metallicity in the CGM (Gutcke et al. 2016)

In this article, depletion times for HI and H₂ (Boselli et al. 2014) and mass-loading factors from CO and CH⁺ data (Cicone et al. 2014; Falgarone et al. 2017) have been added to the list. The NIHAO simulations' ability to reproduce several observational constraints simultaneously gives us confidence that, even if the physics of stellar feedback are not resolved, their effects are described correctly. This confidence is important. Without it, there would be no reason to believe that what we learn from our simulations applies to the Universe.

Many of our results confirm findings that were already known from previous studies (Munshi et al. 2013; Hopkins et al. 2014; Shen et al. 2014; Anglés-Alcázar et al. 2017; Mitchell et al. 2018), but our work delineates a new coherent picture of stellar

feedback as a multiscale and multiphase phenomenon. Stellar feedback operates on all scales from the ISM to several virial radii. Outflows have different components that we must separate for a meaningful comparison with observational data. We summarize our key results proceeding from the smallest scales to the largest ones.

On ISM scales, feedback reduces the efficiency of star formation via two mechanisms: it prevents the condensation of atomic gas into molecular clouds and it moves gas from the cold ISM to a warm/hot phase with $T > 1.5 \times 10^4$ K. The first process, which we call *anticondensation feedback*, is responsible for the long HI's depletion times in dwarf galaxies (Boselli et al. 2014 measures values longer than the Hubble time for galaxies with $M_* \lesssim 10^9 M_\odot$). The specificity of anticondensation feedback is that it does not expel or heat the gas. It simply maintains it to a temperature $T > 1000$ K at which it cannot reach the density required for star formation. In contrast, the warm/hot ISM with $T > 1.5 \times 10^4$ K is not a significant mass reservoir in any galaxies, but even less so in dwarves, where gas vents out as soon as it is heated. Its significance is rather that it is the phase from which winds originate.

On galactic scales, galaxies with $M_{\text{vir}} \sim 3 \times 10^{10} M_\odot$ can expel more than 95 per cent of the accreted baryons but most of this gas (50 per cent–70 per cent) cools very rapidly and falls back in a fountain-like motion on a time-scale of the order of 1 Gyr, which is shorter at higher masses (equation 11). *Fountain feedback* reduces the masses of galaxies by a factor of 2–4 by reducing the time during which gas is available to form stars (in dwarves, gas spends thrice more time in the fountain than it does in the galaxy).

On the scale of the CGM, SN heating offsets the radiative losses and explains the hot CGM's very long cooling time. This *maintenance feedback* is analogous to the one by supermassive black holes in X-ray groups and clusters (e.g. Cattaneo et al. 2009, and references therein).

Expulsion from the halo has probability of 10 per cent–30 per cent per ejection event in massive spirals and dwarves, respectively. Hence, it is a less likely outcome than reaccretion through the fountain. However, after several ejection and reaccretion episodes 20 per cent–70 per cent of the baryons that accrete on to galaxies escape from the virial radius (*ejective feedback*). These baryons seldom come back.

On megaparsec scales, galactic winds can divert the infalling gas while it is still several virial radii away and prevent its accretion on to galaxies. We call this mechanism *pre-emptive feedback*. The ability of SNe to reduce the gas inflow rate through a sphere of radius $r \sim 0.5-1r_{\text{vir}}$ at low masses had already been remarked in cosmological simulations at $z \gtrsim 3$ by Mitchell et al. (2018). However, we have been the first to show that outflows can disrupt inflows and affect the kinematics of gas on scales as large as $6r_{\text{vir}}$. Our work shows that SNe and not photoionization are responsible for the low baryonic fractions of haloes with $M_{\text{vir}} \lesssim 3 \times 10^{10} M_\odot$.

The low M_*/M_{vir} ratios of galaxies are the result of not one but all the above mechanisms put together. Pre-emptive feedback reduces the baryons that can accrete on to a halo with $M_{\text{vir}} \sim 3 \times 10^{10} M_\odot$ to half of the universal baryon fraction. Due to ejective feedback, only 30 per cent of these baryons are retained. Only a quarter of this 30 per cent are in the galaxy because the rest are in the fountain, and, due to anticondensation feedback, the baryons in the galaxy form stars on a very long time-scale (Fig. 10). If an average gas particle spends only ~ 1 Gyr inside the galaxy and the star formation time-scale is ~ 10 Gyr, then only a tenth of the baryons in the galaxy will be converted into stars. Only by multiplying all these factors

do we find $M_*/M_{\text{vir}} \sim 0.004$, which is what we expect from AM for a galaxy with $M_{\text{vir}} \sim 3 \times 10^{10} M_\odot$.

At even lower masses, star formation becomes sporadic. Ultra-dwarves are characterized by occasional bursts of star formation interspersed by long period of quiescence, during which there are not enough SNe for outflows of any consequence. In these galaxies, anticondensation feedback and pre-emptive feedback are the main mechanisms.

Massive spirals accrete nearly 100 per cent of the universal baryon fraction and expel only a small fraction of their baryons. This is due to their deeper gravitational potential wells but also to the appearance of massive hot CGMa that confines winds within the central regions of DM haloes. The fountain is still operating but the mass it processes is at least four times lower, so that the average gas particle spends more time in the galaxy than it does in the fountain; and star formation is much more efficient in massive objects, which have a much higher molecular fraction ($t_{\text{sf}} \sim 1$ Gyr). All these differences combined explain why anticondensation, fountain, ejective, and pre-emptive feedback are much less effective in massive spirals than they are in dwarves. Maintenance is the only process that becomes more important at high masses for the simple reason that low-mass haloes do not contain much hot gas, hence, it makes no difference whether it cools or not. In massive spirals, we also observe the appearance of a hot ISM, which is practically absent in dwarf galaxies.

Our picture of SN feedback is very different from the frequently portrayed one, in which the ejected gas mixes with the hot CGM, so that its reaccretion is determined by the hot CGM's radiative cooling time. We find that only a small fraction of the baryons that accrete on to galaxies are ejected and reaccreted through this route.¹²

A physical picture of galactic winds requires considering their multiphase structure. We propose a model where winds have a hot and a cold component. Blastwaves from SNe form the hot ISM, which vents out of galaxies if it has enough pressure to overcome the depth of the gravitational potential well and the pressure of the surrounding CGM; this is the origin of the hot phase. The cold phase is mainly composed of cold gas that has been entrained by the hot wind. The cold phase is completely dominant at low masses, since there is no need of powerful winds to lift gas from the shallow potential wells of dwarf galaxies. At $M_* \sim 10^{11} M_\odot$, however, the hot component is often dominant. It is only when we separate the cold phase from the hot phase that we are able to reproduce the outflow rates from IRAM (Cicone et al. 2014) and ALMA (Falgarone et al. 2017) over a mass range that spans two orders of magnitude in M_* .

Separating cold and hot outflows is important not only for a correct comparison with observations but also because of their different destinies. The hot component (gas that moves to $r > r_g$ with $T > 1.5 \times 10^4$ K) usually escapes from the halo and never comes back. The cold component falls back on to the disc and forms the fountain. SAMs of galaxy formation could be improved by considering the presence of these two components and by using the NIHAO simulations to calibrate their importance:

$$\begin{aligned} \eta_{\text{gal}} &= 23 \left(v_{\text{vir}} / 67 \text{ km s}^{-1} \right)^{-4.6} \\ \eta_{\text{gal}}^{\text{cold}} &= 23 \left(v_{\text{vir}} / 67 \text{ km s}^{-1} \right)^{-5.6} \end{aligned} \quad (12)$$

¹²The red solid curve in Fig. 18 suggests that this fraction is < 1 per cent, although the real value could be larger if the black solid curve included a contribution from hot atmospheres that have spilled just a little outside r_{vir} .

for $v_{\text{vir}} > 67 \text{ km s}^{-1}$ and

$$\eta_{\text{gal}} = \eta_{\text{gal}}^{\text{cold}} = 23 \left(v_{\text{vir}} / 67 \text{ km s}^{-1} \right)^{-2} \quad (13)$$

for $v_{\text{vir}} < 67 \text{ km s}^{-1}$.

Equations (12) and (13) in conjunction with equation (11) for the cold component's reaccretion time-scale are the equations that we recommend as our reference model. We anticipate, however, that an SAM based on these equations will be successful only if its cooling rate is as low as the one that we measure in the NIHAO simulations, that is, only if it also includes some form of maintenance feedback.

ACKNOWLEDGEMENTS

We thank G. A. Mamon for assistance in computing the stellar masses of M82, NGC 3256, NGC 3628, NGC 253, and NGC 2146. AC would also like to thank C. B. Brook for assistance with PYNBODY. We thank N. Bouché for his useful comments. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (<http://www.lrz.de>) and the High Performance Computing resources at New York University Abu Dhabi.

REFERENCES

- Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Quataert E., Murray N., 2017, *MNRAS*, 470, 4698
- Baldry I. K. et al., 2012, *MNRAS*, 421, 621
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
- Beirão P. et al., 2015, *MNRAS*, 451, 2640
- Bennett C. L. et al., 2013, *ApJS*, 208, 20
- Bertschinger E., 2011, *Philos. Trans. R. Soc. Lond. Ser. A*, 369, 4947
- Binney J., 1977, *ApJ*, 215, 483
- Birzan L., Rafferty D. A., McNamara B. R., Wise M. W., Nulsen P. E. J., 2004, *ApJ*, 607, 800
- Boselli A., Lequeux J., Gavazzi G., 2002, *A&A*, 384, 33
- Boselli A., Cortese L., Boquien M., Boissier S., Catinella B., Lagos C., Saintonge A., 2014, *A&A*, 564, A66
- Bouché N., Hohensee W., Vargas R., Kacprzak G. G., Martin C. L., Cooke J., Churchill C. W., 2012, *MNRAS*, 426, 801
- Brook C. B., Stinson G., Gibson B. K., Wadsley J., Quinn T., 2012, *MNRAS*, 424, 1275
- Brook C. B., Stinson G., Gibson B. K., Shen S., Macciò A. V., Obreja A., Wadsley J., Quinn T., 2014, *MNRAS*, 443, 3809
- Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80
- Cattaneo A. et al., 2009, *Nature*, 460, 213
- Cattaneo A. et al., 2017, *MNRAS*, 471, 1401
- Chabrier G., 2003, *ApJ*, 586, L133
- Christensen C. R., Davé R., Governato F., Pontzen A., Brooks A., Munshi F., Quinn T., Wadsley J., 2016, *ApJ*, 824, 57
- Cicone C. et al., 2014, *A&A*, 562, A21
- Crain R. A. et al., 2015, *MNRAS*, 450, 1937
- Daddi E. et al., 2007, *ApJ*, 670, 156
- de Vaucouleurs G., de Vaucouleurs A., Corwin H. G., Jr., Buta R. J., Paturel G., Fouqué P., 1991, in *Third Reference Catalogue of Bright Galaxies. Volume I: Explanations and references. Volume II: Data for galaxies between 0^h and 12^h . Volume III: Data for galaxies between 12^h and 24^h* , Springer, New York
- Dekel A., Birnboim Y., 2006, *MNRAS*, 368, 2
- Dekel A., Silk J., 1986, *ApJ*, 303, 39
- Dutton A. A. et al., 2017, *MNRAS*, 467, 4937
- Dutton A. A., Macciò A. V., 2014, *MNRAS*, 441, 3359
- Falgarone E. et al., 2017, *Nature*, 548, 430
- Fioc M., Rocca-Volmerange B., 1997, *A&A*, 326, 950
- Geha M., Guhathakurta P., Rich R. M., Cooper M. C., 2006, *AJ*, 131, 332
- Guedes J., Callegari S., Madau P., Mayer L., 2011, *ApJ*, 742, 76
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Gutcke T. A., Stinson G. S., Macciò A. V., Wang L., Dutton A. A., 2016, *MNRAS*, 464, 2796
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, *MNRAS*, 431, 3373
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, 451, 2663
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581
- Kacprzak G. G. et al., 2014, *ApJL*, 792, L12
- Keller B. W., Wadsley J., Benincasa S. M., Couchman H. M. P., 2014, *MNRAS*, 442, 3013
- Kennicutt R. C., Jr., 1998, *ApJ*, 498, 541
- Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *MNRAS*, 363, 2
- Knollmann S. R., Knebe A., 2009, *ApJS*, 182, 608
- Lacey C. G. et al., 2016, *MNRAS*, 462, 3854
- Leauthaud A. et al., 2012, *ApJ*, 744, 159
- Lehnert M. D., Heckman T. M., Weaver K. A., 1999, *ApJ*, 523, 575
- Leroy A. K. et al., 2015, *ApJ*, 814, 83
- Macciò A. V., Udrescu S. M., Dutton A. A., Obreja A., Wang L., Stinson G. R., Kang X., 2016, *MNRAS*, 463, L69
- McKee C. F., Ostriker J. P., 1977, *ApJ*, 218, 148
- McKeith C. D., Greve A., Downes D., Prada F., 1995, *A&A*, 293, 703
- Mitchell P. D., Blaizot J., Devriendt J., Kimm T., Michel-Dansac L., Rosdahl J., Slyz A., 2018, *MNRAS*, 474, 4279
- Mitra S., Davé R., Finlator K., 2015, *MNRAS*, 452, 1184
- Moster B. P., Naab T., White S. D. M., 2013, *MNRAS*, 428, 3121
- Moustakas J. et al., 2013, *ApJ*, 767, 50
- Munshi F. et al., 2013, *ApJ*, 766, 56
- Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, *MNRAS*, 454, 2691
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Nelson D., Genel S., Vogelsberger M., Springel V., Sijacki D., Torrey P., Hernquist L., 2015, *MNRAS*, 448, 59
- Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, *MNRAS*, 406, 2325
- Papastergis E., Cattaneo A., Huang S., Giovanelli R., Haynes M. P., 2012, *ApJ*, 759, 138
- Penzo C., Macciò A. V., Casarini L., Stinson G. S., Wadsley J., 2014, *MNRAS*, 442, 176
- Pereira-Santaella M. et al., 2016, *A&A*, 594, A81
- Persic M., Salucci P., Stel F., 1996, *MNRAS*, 281, 27
- Pizagno J. et al., 2007, *AJ*, 134, 945
- Planck Collaboration XX et al., 2014, *A&A*, 571, A20
- Pontzen A., Roškar R., Stinson G., Woods R., 2013, in *pynbody: N-Body/SPH analysis for python, Astrophysics Source Code Library, ascl:1305.002*
- Rafferty D. A., McNamara B. R., Nulsen P. E. J., Wise M. W., 2006, *ApJ*, 652, 216
- Rees M. J., Ostriker J. P., 1977, *MNRAS*, 179, 541
- Ritchie B. W., Thomas P. A., 2001, *MNRAS*, 323, 743
- Saitoh T. R., Makino J., 2009, *ApJL*, 697, L99
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Schroetter I. et al., 2016, *ApJ*, 833, 39
- Schroetter I., Bouché N., Péroux C., Murphy M. T., Contini T., Finley H., 2015, *ApJ*, 804, 83
- Shen S., Wadsley J., Stinson G., 2010, *MNRAS*, 407, 1581
- Shen S., Madau P., Conroy C., Governato F., Mayer L., 2014, *ApJ*, 792, 99
- Silk J., 1977, *ApJ*, 211, 638
- Silk J., 2003, *MNRAS*, 343, 249
- Springob C. M., Masters K. L., Haynes M. P., Giovanelli R., Marinoni C., 2007, *ApJS*, 172, 599
- Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, *MNRAS*, 373, 1074
- Stinson G. S., Brook C., Macciò A. V., Wadsley J., Quinn T. R., Couchman H. M. P., 2013, *MNRAS*, 428, 129

- Strickland D. K., Heckman T. M., 2009, *ApJ*, 697, 2030
Sutherland R. S., Dopita M. A., 1993, *ApJS*, 88, 253
Tasca L. A. M. et al., 2015, *A&A*, 581, A54
Tollet É., Cattaneo A., Mamon G. A., Moutard T., van den Bosch F. C., 2017, *MNRAS*, 471, 4170
Tomczak A. R. et al., 2014, *ApJ*, 783, 85
Torres G., Andersen J., Giménez A., 2010, *A&AR*, 18, 67
Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
Wadsley J. W., Stadel J., Quinn T., 2004, *New Astron.*, 9, 137
Wadsley J. W., Veeravalli G., Couchman H. M. P., 2008, *MNRAS*, 387, 427
Wadsley J. W., Keller B. W., Quinn T. R., 2017, *MNRAS*, 471, 2357
Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, *MNRAS*, 454, 83
Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Gutcke T., Kang X., 2017, *MNRAS*, 466, 4858
Yang X., Mo H. J., van den Bosch F. C., 2009, *ApJ*, 693, 830

This paper has been typeset from a \TeX/L\AA\TeX file prepared by the author.