Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRASL **442,** L105–L109 (2014) doi:10.1093/mnrasl/slu066

Overconsumption, outflows and the quenching of satellite galaxies

Sean L. McGee,[1⋆] Richard G. Bower[2] and Michael L. Balogh[1,3]

[1]*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[2]*Department of Physics, University of Durham, Durham DH1 3LE, UK*
[3]*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada*

Accepted 2014 May 1. Received 2014 April 23; in original form 2014 February 25

ABSTRACT

The baryon cycle of galaxies is a dynamic process involving the intake, consumption and ejection of vast quantities of gas. In contrast, the conventional picture of satellite galaxies has them methodically turning a large gas reservoir into stars until this reservoir is forcibly removed due to external ram pressure. This picture needs revision. Our modern understanding of the baryon cycle suggests that in some regimes the simple interruption of the fresh gas supply may quench satellite galaxies long before stripping events occur, a process we call overconsumption. We compile measurements from the literature of observed satellite quenching times at a range of redshifts to determine if satellites are principally quenched through orbit-based gas stripping events – either direct stripping of the disc (ram pressure stripping) or the extended gas halo (strangulation) – or from internally driven star formation outflows via overconsumption. These time-scales show significant deviations from the evolution expected for gas stripping mechanisms and suggest that either ram pressure stripping is much more efficient at high redshift, or that secular outflows quench satellites before orbit-based stripping occurs. Given the strong redshift evolution of star formation rates, at high redshift even moderate outflow rates will lead to extremely short delay times with the expectation that high-redshift ($z > 1.5$) satellites will be quenched almost immediately following the cessation of cosmological inflow. Observations of high-redshift satellites give an indirect but sensitive measure of the outflow rate, with current measurements suggesting that outflows are no larger than 2.5 times the star formation rate for galaxies with a stellar mass of $10^{10.5}$ $M_{\odot}$.

Key words: galaxies: clusters: general – galaxies: evolution – galaxies: formation – galaxies: groups: general.

1 INTRODUCTION

The conventional picture of the interaction between galaxies and their environments is based on the idea that galaxies enter dense environments with a reservoir of gas. The removal of this reservoir, either from the stellar disc (ram pressure stripping), or the galaxy's halo (strangulation), leads to a decline in the star formation rate (SFR). This view is out of date with our understanding of the dynamic baryon cycle in galaxies. Cosmological simulations, as well as semi-analytic models, show that galaxies grow as a result of gas infall from surrounding filaments. At least at high redshift, rapid star formation is fed by a supply of fresh, infalling gas which dominates over the consumption of any reservoir (e.g. Papovich et al. 2011). Taking this into account leads to a significant update to our picture of how galaxies interact with their environment. If this scenario is correct, simply removing the supply of fresh material will have a dramatic effect on the properties of satellite galaxies, without the need for gas stripping processes. In this Letter, we use a compilation of observed satellite star formation time-scales at a range of redshifts, to examine the mechanisms which quench satellites and the insights this gives to the baryon cycle of all galaxies.

In recent years, many authors have combined direct observations of satellite quenched fractions with cosmological infall rates of galaxies to estimate the time it takes a satellite to quench; the so-called quenching time (e.g. McGee et al. 2011; De Lucia et al. 2012; Wetzel et al. 2013). This is interpreted as the length of time after 'accretion', where accretion is loosely associated with the time, a galaxy is within the virialized region of some larger structure, until the galaxy is completely quenched. At low redshift, this time-scale is long (several Gyr; Wetzel et al. 2013), leading to the suggestion that gentle stripping of the gas reservoir is the mechanism responsible. However, such a picture naturally leads to most satellite galaxies having relatively paltry SFRs as their gas is dwindled away.

In direct contrast, observational studies of satellites have been hampered by the lack of such 'smoking guns' of galaxy transformation. Indeed, many recent studies have shown that those satellites which are forming stars are doing so at a rate indistinguishable from central galaxies of similar mass (e.g. Peng et al. 2010; McGee et al. 2011; Muzzin et al. 2012; Wetzel, Tinker & Conroy 2012).

⋆ E-mail: mcgee@strw.leidenuniv.nl

To parametrize this behaviour, Wetzel et al. (2013) introduced the notion of a 'delay time' – that is, the length of time after infall for which a satellite galaxy appears to track the SFR of an analogous central galaxy.[1] In context of a rapid baryon cycle, the length of this delay time holds clues for our understanding of galaxy formation as it is expected that the fresh gas supply is also cutoff near infall. Dark matter simulations show that the accretion of mass into a subhalo stops when it reaches 1.8 $R_{\rm vir}$ (as defined in Bryan & Norman 1998) of a larger halo (Behroozi et al. 2014). This limits the reservoir of gas which can be used for future star formation of the satellite, even in the absence of stripping.

The final relevant time-scale of satellite quenching is a 'fading time'. This is equal to the quenching time minus the delay time, and is interpreted physically as the time it takes a satellite to be quenched once its SFR deviates from that of an analogous, central galaxy. This phase is short (<1 Gyr) and is reflected in the relative lack of 'green' or intermediate star-forming galaxies.

In this Letter, we compile measurements of the quenching and delay times of satellites at a range of redshifts to determine if 'orbit-based' or 'outflow-based' models dominate the quenching. Further, we can directly put limits on the outflow rates allowed by such delay times. We finish by exploring the difficulty of achieving long delay times in standard models of galaxy formation. In this Letter, we adopt a Λ cold dark matter cosmology with the parameters: $\Omega_{\rm m} = 0.27$, $\Omega_\Lambda = 0.73$ and $h = H_0/(100\,{\rm km\,s^{-1}\,Mpc^{-1}}) = 0.70$. All stellar masses assume a Chabrier (2003) initial mass function (IMF).

2 OBSERVED SATELLITE QUENCHING TIME-SCALES

The best current measurement comes from the analysis of Wetzel et al. (2013) from a low-redshift ($z \sim 0.05$) Sloan Digital Sky Survey sample of galaxy groups, where the activity level of the galaxies is determined from an Hα-based spectroscopic measurement. Wetzel et al. have carefully tracked the infall history of satellites at this epoch, accounting for the evolution in infalling field SFRs, to find that a satellite galaxy with a stellar mass of $10^{10.5}$ M$_\odot$ has a total quenching time of 4.4 $\pm$ 0.4 Gyr. They find that 3.8 $\pm$ 0.4 Gyr of this time is due to the 'delay time', while the remaining ∼0.6 Gyr of the quenching time is the 'fading' stage. Wetzel et al. (2013) define the moment a galaxy becomes a 'satellite' as the time the subhalo crosses 200 times the matter background density ($R_{200,\rm b}$), not the critical density ($R_{200,\rm c}$). This $R_{200,\rm b}$ definition is similar to the $R_{\rm vir}$ used in Behroozi et al. (2014). These densities evolve differently with redshift, so we consistently assume the Wetzel et al. definition and make corrections when the observational results were presented differently. At $z = 0.0$, this matter density radius is 1.92 times larger than the critical density radius, although by $z = 1.0$ it is only 1.15 times larger.

Using data from the Group Environment Evolution Collaboration 2 sample of spectroscopic galaxy groups at $z \sim 0.9$, Mok et al. (2014) presented an estimate of the quenching and delay time-scales by examining the relative fractions of 'star-forming', 'intermediate' and 'quiescent' satellites as defined by their locations in a colour–colour plane designed to separate the contributions from age and dust. Using an infall model and parametrization of the star formation histories of central galaxies, these authors concluded that these $z \sim 0.9$ satellites are best fitted by a total quenching time of 1.05 $\pm$ 0.25 Gyr, made up of a 0.8 $\pm$ 0.25 Gyr delay time and the remaining ∼0.25 Gyr due to the fading time. This time-scale applies to $10^{10.5}$ M$_\odot$ galaxies, but is not seen to be a strong function of stellar mass for this sample.

From a recent comparison between the phase space distribution of post-starburst galaxies in $z \sim 1$ galaxy clusters from the Gemini CLuster Astrophysics Spectroscopic Survey sample and the phase space of subhaloes accreted into cosmological-zoom simulations of galaxy clusters, Muzzin et al. (2014) find that the total time-scale for quenching is $\sim 1 \pm 0.25$ Gyr. The characteristic phase space features of these post-starburst galaxies can be best represented if quenching occurs quickly (0.1–0.5 Gyr) following the first passage of 0.25–0.5 $R_{200,\rm c}$. This is consistent with an analysis of the spectral features of the post-starbursts, selected to have low $D_n(4000)$ and a lack of [O II] emission, which suggests that they have a short fading time (∼ 0.1–0.5 Gyr). This implies that the delay time is the travel time from $R_{200,\rm b}$ to this position. Directly from this simulation, this delay time is then 0.45±0.15 Gyr depending on which inner distance the quenching occurs at. While the Muzzin et al. time-scale is not presented in bins of stellar mass, they do state that most of the relevant galaxies are in the range of $10^{10.5}$ $M_{\rm stel}$, and thus should be comparable to our previous time-scales.

We also add a quenching time derived from the *GALEX* -based SFRs of galaxy groups at z ∼0.4 from McGee et al. (2011). We can define the excess quenching between groups and the field in a manner similar to van den Bosch et al. (2008) as $\frac{f_{\rm q,grp} - f_{\rm q,fld}}{1 - f_{\rm q,fld}}$, where $f_{\rm q,grp}$ and $f_{\rm q,fld}$ are the quenched fraction of the groups and the field, respectively. In $10^{10.5}$ M$_\odot$ galaxies, this excess quenching is 0.32 $\pm$ 0.1. By comparison with the accretion histories as compiled in McGee et al. (2009), we find that at $z = 0.4$, 30 per cent of galaxies became satellites more than 4.0± 0.6 Gyr ago. Although these authors do not explicitly measure the fading time, for our purposes we also assume that it is 0.5 Gyr. We caution that even minor differences in the criteria to demarcate between star-forming and passive can lead to large differences within different surveys. In the future, a more comprehensive and homogeneous analysis of the existing data will be useful.

3 ANALYSIS

3.1 Orbit-based or outflow-based quenching?

The time it takes a galaxy to proceed from the virial radius of a galaxy cluster to a specific location (e.g. perigee) in its orbit depends on redshift. Perhaps the simplest hypothesis is that the efficiency of gas stripping depends only on the orbit. That is, the same satellite will be stripped of the same fraction of gas at perigee, in the same orbit, regardless of whether this occurs at $z = 0$ or $z = 1$. In this case, the quenching time-scale of galaxies from 'orbit-based' quenching mechanisms will evolve as the ratio of the inverse densities, $\propto (1 + z)^{-3/2}$.[2]

[1] The existence of this delay time is still a matter of some debate as several authors have found that star-forming galaxies have lower specific star formation rates (sSFRs) in satellites than in centrals (e.g. Patel et al. 2009; Vulcani et al. 2010; Haines et al. 2013). However, some of this discrepancy results from the definition of star-forming galaxies, where authors who have more inclusive definitions find the discrepancy. Nevertheless, the existence of a tail of low-SFR galaxies does not affect the overall observation that most satellites are forming stars at normal rates for a significant time.

[2] The crossing time of a cluster is the ratio of the radius, R, and the velocity, V. For a cluster in virial equilibrium the cluster mass, M, scales as RV^2 and for a spherical cluster as $R^3\rho$. So, the dynamical time, R/V, scales as $\rho^{-1/2}$, while the density, ρ, scales as $(1+z)^3$.

In contrast, an outflow-based quenching mechanism should have a time-scale which scales as the SFR. Once the cosmological accretion of gas is halted from penetrating into the sphere of influence of a satellite galaxy, the time-scale on which the galaxy quenches is equal to the total gas available at this point divided by the 'gas consumption' rate. While this gas consumption rate is usually taken to be equal to the SFR of the galaxy, this ignores the outflows of gas associated with this star formation. It is partly for this reason that satellite galaxies in semi-analytic models are quenched more quickly than the SFR alone would suggest. In this scenario, satellites prematurely halt their growth because they continue vigorous star formation, and drive associated outflows, despite a significant restriction on their available resources. We therefore borrow a term from the field of ecology and refer to this quenching mechanism as 'overconsumption'.

For our purposes in this section, the exact outflow rate is not important if it is invariant in redshift for a given SFR and stellar mass ($M_{\rm stel}$). In that case, any outflow-based quenching would require the quenching time to scale with the SFR, or the specific star formation rate (sSFR = SFR/$M_{\rm stel}$). Several authors have presented fitting formula to the observed star-forming 'sequence' of galaxies over cosmic time. In particular, Peng et al. (2010) found that

$$\mathrm{sSFR} = 2.5 \left(\frac{M_{\rm stel}}{10^{10}\,\mathrm{M}_\odot}\right)^{\beta} \left(\frac{t}{3.5\ \mathrm{Gyr}}\right)^{-2.2}, \quad (1)$$

where t is the age of the Universe and β describes the tilt of the sSFR relation, which we take to be 0. Similarly, Whitaker et al. (2012) parametrized the SFR as

$$\log(\mathrm{SFR}) = \alpha(z)(\log M_{\rm stel} - 10.5) + \beta(z), \quad (2)$$

where $\alpha(z) = 0.70 - 0.13z$ and $\beta(z) = 0.38 + 1.14z - 0.19z^2$. We include these two relations as a measure of the uncertainty in such evolutionary measurements of the sSFR, and note that such relations rely heavily on the assumed functional form at very low and very high redshift where the fit is not well constrained by the data (e.g. Behroozi, Wechsler & Conroy 2013).

In Fig. 1, we show the evolution of quenching time expected from orbit-based and outflow-based quenching (using either the Peng et al. or Whitaker et al. SFRs with $M_{\rm stel} = 10^{10.5}\,\mathrm{M}_\odot$) when normalized to a quenching time of 4.4 Gyr at $z = 0.05$. These are shown along with the observed quenching times described in Section 2. While the lower redshift point is roughly consistent with the dynamical time evolution, the $z \sim 1$ quenching times seem to be lower than expected. This could be due either to an onset of outflow-based quenching or an enhanced stripping efficiency at high redshift. We note that Tinker & Wetzel (2010) and Tinker et al. (2013) have concluded that their halo model/clustering measurements are consistent with dynamical evolution, although they do not compare with SFR-derived time-scales. In the next section, we will determine the outflow rates required to explain these quenching times.

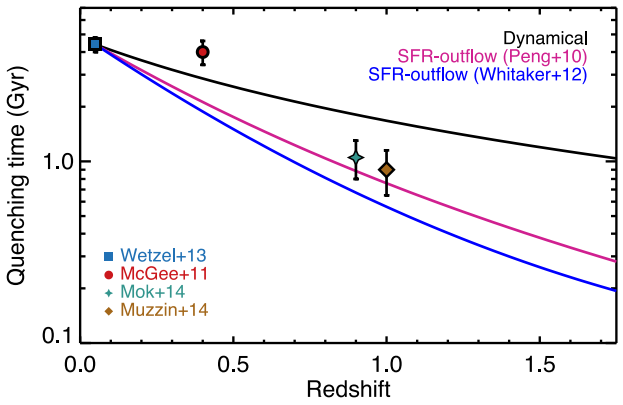

Figure 1. The evolution of satellite quenching time with redshift. The lines represent the expectation from an 'orbit-based' or 'outflow-based' quenching mechanism. The data points represent the observed quenching times of Wetzel et al. (2013), McGee et al. (2011), Mok et al. (2014) and Muzzin et al. (2014) as described in Section 2.

3.2 Limits on the outflow rate

As discussed in the introduction, one of the most curious observed aspects of satellite galaxies is that those which are forming stars appear to have similar SFRs to central galaxies of similar mass – requiring a 'delay time'. Given the halting of cosmological accretion, the largely unchanging SFR during the delay time must result from an unchanged fuelling rate of the cold gas disc from the halo's gas reservoir, and will result in an unchanged outflow rate driven by this star formation. Thus, even if the delay time is established by the onset of orbit-based quenching, there must be a sufficient reservoir at satellite infall to maintain this fuelling throughout. Given the delay time, we can determine the maximum allowed outflow rate from basic baryon accounting measures.

The inflow and outflow of gas from a galaxy halo can be parametrized as

$$\dot{M}_{\rm in} + \dot{M}_{\rm res} = \dot{M}_{\rm stel}(1 - R) + \dot{M}_{\rm out}, \quad (3)$$

where $\dot{M}_x$ is the rate of gas accretion due to cosmological infall (x = in), the rate of change of the gas reservoir of the system (x = res), the SFR (x = stel) and the rate of outflows from feedback (x = out). R is fraction of gas returned to the interstellar medium through supernova, stellar winds, etc., and we assume that this is done instantaneously. This basic equation has been the basis of galaxy formation models for some time (e.g. Cole et al. 2000), and has recently been recast as the 'equilibrium condition' by several authors (e.g. Bouché et al. 2010; Davé, Finlator & Oppenheimer 2012; Lilly et al. 2013), although the precise definition varies depending on the definition of the system (e.g. as either the galaxy alone, or the galaxy and its extended halo) and whether a 'reservoir' exists. Note that a non-zero delay time immediately suggests that a sizeable reservoir must exist upon satellite infall.

Assuming that cosmological accretion is halted on infall, then $\dot{M}_{\rm in} = 0$. As noted previously, there are simulations which suggest this happens to the dark matter prior to $R_{\rm vir}$ and, in such cases, we would overestimate the maximum outflow rate. The maximum time that the SFR can be maintained at a constant rate is

$$T_{\rm delay} = \frac{M_{\rm res}}{\dot{M}_{\rm res}}. \quad (4)$$

Taking account of the possible locations of the halo's baryon content at the point of infall, $M_{\rm res}$ can be given from a conservation equation

$$M_{\rm res} = M_{\rm baryon} - M_{\rm cold} - M_{\rm stel} - M_{\rm ej} - M_{\rm strip}, \quad (5)$$

where $M_{\rm baryon}$ is the total mass of baryons that *belong* to the system, and can be taken as the universal fraction of baryons multiplied by the satellite halo mass ($f_{\rm baryon}{}^{*}M_{\rm halo}$). $M_{\rm ej}$ is the mass of baryons that have been ejected from this system prior to infall. In our discussion, we will assume that baryons ejected are lost from the system, with no re-accretion. By this definition, the integral of $\dot{M}_{\rm out}$ from its formation time to infall is equal to $M_{\rm ej}$. This is largely just a definition of our terms, and therefore defines the outflow rate to be the rate at which gas is expelled from the system forever. $M_{\rm strip}$ is the mass

of gas which could be stripped during the lifetime of the quenching phase, and can be held to zero to find the maximum possible outflow rate. $M_{\rm cold}$ is the cold gas currently locked up in the galaxy. It is likely that this mass must be approximately constant throughout the delay time in order to produce an approximately constant SFR according to the Kennicutt–Schmidt law (Schmidt 1959; Kennicutt 1998). Perhaps, it is the eventual depletion of this gas that results in the 'fading time'.

In the case of outflows produced from star formation, either directly through radiation pressure or from subsequent supernova explosions, the outflow rate is related to the SFR and is often parametrized as a mass-loading factor, η ($\equiv \dot{M}_{\rm out}/\dot{M}_{\rm stel}$). Using the equilibrium equation (3), along with the conservation equation (5) and the definition of the mass-loading factor, the delay time is

$$T_{\rm delay} = \frac{M_{\rm baryon} - M_{\rm cold} - M_{\rm stel} - M_{\rm ej} - M_{\rm strip}}{M_{\rm stel}(1 - R + \eta){\rm sSFR}}, \quad (6)$$

where we have defined the specific star formation rate (sSFR $= \dot{M}_{\rm stel}/M_{\rm stel}$). For the purposes of parametrizing our knowledge of the baryonic components, it is convenient to define each component with respect to the halo mass at infall, so that f_x is $M_x/M_{\rm halo}$. R, the returned gas fraction, is usually of the order of ∼0.25–0.4, depending on the choice of IMF and the time since the stellar population was born. We assume a value relevant for a Chabrier IMF with a 2 Gyr population ($R = 0.4$).

We can now estimate each of these components. In our assumed cosmology, $f_{\rm baryon} = 0.17$. There have been many attempts to quantify $f_{\rm stel}$, largely through halo abundance matching techniques. It has been shown that this fraction is constant with redshift, assuming a universal IMF, although it does strongly depend on halo mass (e.g. Guo et al. 2010; Moster et al. 2010; Behroozi et al. 2013; McGee, Goto & Balogh 2014). For our canonical $10^{10.5}$ M$_\odot$ stellar mass galaxy, using the functional form presented by Moster et al. (2010), we find that this stellar fraction is 0.03. Although we adopt $f_{\rm stel} = 0.03$, it is worth noting that Behroozi et al. (2013) find similar stellar fractions in the range 0.024–0.028 to $z \sim 3$ at this stellar mass. Carilli & Walter (2013) have compiled existing measurements of the H_2 content of strongly star-forming galaxies at a range of redshifts and found that $f_{\rm cold} = 0.1 * (1+z)^2 f_{\rm stel}$. As the H_2 is expected to correlate well with the SFR, this is an appropriate value for the cold gas content, rather than the sum of molecular and atomic hydrogen gas. In our parametrization, atomic hydrogen will be part of the gas reservoir. If we assume that the ejected gas is a result of past outflows from SFRs, then $f_{\rm ej} = \eta f_{\rm stel}$ (1+R), where the R is needed to recover the total past SFR. This model of past gas ejection is clearly an approximation, as earlier star formation occurred as the halo was forming and thus η is unlikely to be constant throughout. Indeed, given the lower potential wells, and the need to widely eject metals into the intergalactic medium to match observation, this η may be an underestimate of the amount of ejected material. As well, this assumes that the star formation driven outflows dominate the ejected gas budget.

Using these approximations, the delay time is

$$T_{\rm delay} = \frac{f_{\rm baryon} - f_{\rm stel}(0.1 * (1+z)^2 + \eta(1+R) + 1) - f_{\rm strip}}{f_{\rm stel}(1 - R + \eta){\rm sSFR}}. \quad (7)$$

We note that equation (7) has only one free parameter, η, and so can be completely characterized by a measure of the satellite delay time. In Fig. 2, we show the maximum delay times (assuming $f_{\rm strip} = 0$) as a function of redshift for $\eta = 1$, 2, 2.5, 3. We have also plotted the delay times from the observations, and a dynamical

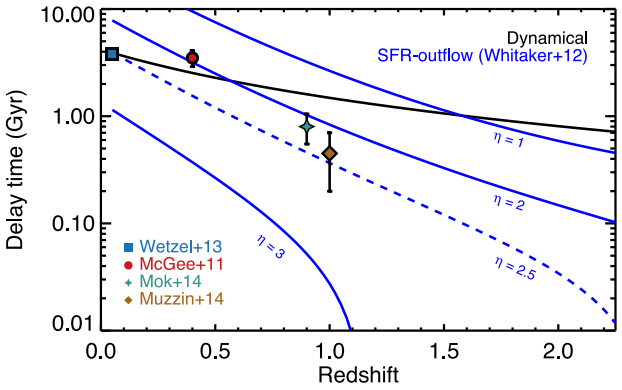

Figure 2. The maximum possible delay times inferred as a function of redshift for a given mass-loading factor ($\eta = 1$, 2, 2.5, 3) and assuming a Whitaker et al. (2012) star formation history of star-forming galaxies. The dynamical time-scale is also shown.

scaling which fits the delay time at $z = 0.05$. Note that the scaling with any particular η is much steeper than the dynamical scaling because the evolution of the observed sSFR is very rapid. The turnover in delay time at $\eta = 3$ and $z > 1.0$, or $\eta = 2.5$ and $z > 2.0$, results from the need to maintain ever larger fractions of cold gas in the disc, while past outflows caused by the high η make this impossible. The incompatible slope of the SFR and that implied by the dynamical time is a long-standing problem for models, and may suggest an underlying change in the efficiency of star formation, either through varying wind re-accretion or small-scale cloud conditions.

Interestingly, in Fig. 2, we see that the $z \sim 1$ observations are consistent with a maximum mass-loading of 2.5. Observations of strongly star-forming galaxies suggest that the mass-loading is of the order of unity (Heckman et al. 2000; Martin 2005), although the measurements are difficult to interpret. While the ultimate fate of this outflowing gas is unclear, some simulations suggest that up to a third of the gas is re-accreted (Oppenheimer & Davé 2008), although this does not necessarily apply in satellites. In semi-analytic models, which are designed to reproduce the luminosity/mass function, the mass-loading factors are typically much greater than 1, and are a strong function of halo mass (e.g. Guo et al. 2011; Bower, Benson & Crain 2012). Although much of this gas may also be re-accreted, it is again not clear if this is the case in satellites (Font et al. 2008). However, note that with even moderate mass-loading of 1, the expected delay time is shorter than expected from a scaling of the dynamical time in $z > 1.5$ satellite galaxies. This is a potentially new quenching mechanism for satellite galaxies, and calls for a systematic observational study of such galaxies.

3.3 Implications for the cooling model

In the previous section, we have implicitly assumed that the only necessary condition to sustain star formation at a constant rate is sufficient gas in the reservoir. However, such a constant SFR also requires that the cooling rate of gas from the reservoir is unchanged throughout the delay time. In a common model of cooling gas, the bulk of the gas reservoir is in the form of a hot halo, with a temperature that is a result of shock heating to the virial temperature upon accretion. This gas cools, largely through bremsstrahlung, and the rate depends on the square of the gas density. In simple models, as the gas near the central, dense region cools, the gas will respond to the flow quickly, so that the gas density profile shape will be

driven by the dark matter potential. This seems like the best case scenario, as it is expected that the hot gas may build up outside the cooling radius (Fabian & Nulsen 1977). Thus, while the profile shape is constant, the decline in the total gas content will cause the central density to decline. The cooling rate is then directly linked to the amount of the gas in the reservoir. As the hot gas reservoir begins to drain, the cooling rate will go down, the subsequent amount of cold gas will decline and thus, star formation will also decline.

One solution to maintain uniform cooling rates is to change the shape of the gas profile as it is drained. It is possible that the immersion of a satellite galaxy into the hot, dense intracluster medium (ICM) will compress the hot halo. However, given the need to match the compression rate with the cooling rate regardless of the exact density of the ICM, this suggests a fine tuning problem. Unfortunately, it is very difficult to observe X-ray haloes of galaxies, and thus to determine this directly. One alternative to such a fine tuning is that the accreted gas is never shock heated to the virial temperature, and simply adds to the atomic gas reservoir. The thermal history of gas in galaxies has long been studied (White & Rees 1978), and is still debated (e.g. Kereš et al. 2005; Dekel & Birnboim 2006). Interestingly, though, this may hint that the cooling rate is governed by the interplay between atomic and molecular gas rather than the hot gas cooling. As we have mentioned, our definition of the gas 'reservoir' includes the H I content of gas. In principle, if this is the primary gas reservoir, we would expect to see a decrease in the H I content for satellite galaxies which are strongly star-forming in relation to that seen in central galaxies. In a future paper, we will directly measure the H I content of star-forming satellites for indications that this is being drained, and thus place stronger constraints on the cooling mode.

4 CONCLUSIONS

We have compiled measurements of the quenching and delay times of satellite galaxies as a function of redshift from the literature. We have used these measurements, in combination with a simple model for the baryon cycle in galaxies, to arrive at the following conclusions:

(i) While our compilation of measurements may be subject to systematic effects, the evolution of quenching time-scales is consistent with the scaling expected if the quenching was driven not by 'orbit-based' stripping events like ram pressure stripping or strangulation but rather by secular outflows in the light of a halt of cosmological accretion (overconsumption). Alternatively, it may be that 'orbit-based' stripping is significantly more effective at high redshift than a dynamical scaling would suggest.

(ii) The delay time of satellite galaxies places an upper limit on the rate of outflow-driven gas, or the mass-loading, which from current data is $\eta = 2.5$ in $M_{\rm stel} = 10^{10.5}\,{\rm M}_{\odot}$ galaxies. Further measurements at higher redshift and lower masses will put even tighter constraints on the outflow rate.

(iii) The delay time requires an uninterrupted rate of cooling gas from the gas reservoir. This presents a challenge to typical models of cooling gas for which the cooling rate is strongly dependent on the density of the reservoir, and thus declines quickly as the reservoir drains.

(iv) We predict that at $z > 1.5$ satellite galaxies are quenched through secular outflow processes instead of external stripping events. Future observations of such galaxies will put constraints on the outflow rates of all star-forming galaxies and probe this unique quenching mechanism.

ACKNOWLEDGEMENTS

SLM acknowledges support from an NWO grant to George Miley. RGB is supported by STFC grant ST/I001573/1. MLB acknowledges support from an NSERC Discovery grant, and from NOVA and NWO visitor grants that supported his sabbatical visit to Leiden Observatory, where this work was undertaken.

REFERENCES

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Behroozi P. S., Wechsler R. H., Lu Y., Hahn O., Busha M. T., Klypin A., Primack J. R., 2014, ApJ, 787, 156
Bouché N. et al., 2010, ApJ, 718, 1001
Bower R. G., Benson A. J., Crain R. A., 2012, MNRAS, 422, 2816
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Carilli C. L., Walter F., 2013, ARA&A, 51, 105
Chabrier G., 2003, PASP, 115, 763
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98
De Lucia G., Weinmann S., Poggianti B. M., Aragón-Salamanca A., Zaritsky D., 2012, MNRAS, 423, 1277
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Fabian A. C., Nulsen P. E. J., 1977, MNRAS, 180, 479
Font A. S. et al., 2008, MNRAS, 389, 1619
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Guo Q. et al., 2011, MNRAS, 413, 101
Haines C. P. et al., 2013, ApJ, 775, 126
Heckman T. M., Lehnert M. D., Strickland D. K., Armus L., 2000, ApJS, 129, 493
Kennicutt R. C., Jr1998, ApJ, 498, 541
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119
McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, MNRAS, 400, 937
McGee S. L., Balogh M. L., Wilman D. J., Bower R. G., Mulchaey J. S., Parker L. C., Oemler A., 2011, MNRAS, 413, 996
McGee S. L., Goto R., Balogh M. L., 2014, MNRAS, 438, 3188
Martin C. L., 2005, ApJ, 621, 227
Mok A. et al., 2014, MNRAS, 438, 3070
Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903
Muzzin A. et al., 2012, ApJ, 746, 188
Muzzin A. et al., 2014, MNRAS, 438, 3070
Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577
Papovich C., Finkelstein S. L., Ferguson H. C., Lotz J. M., Giavalisco M., 2011, MNRAS, 412, 1123
Patel S. G., Holden B. P., Kelson D. D., Illingworth G. D., Franx M., 2009, ApJ, 705, L67
Peng Y. et al., 2010, ApJ, 721, 193
Schmidt M., 1959, ApJ, 129, 243
Tinker J. L., Wetzel A. R., 2010, ApJ, 719, 88
Tinker J. L., Leauthaud A., Bundy K., George M. R., Behroozi P., Massey R., Rhodes J., Wechsler R. H., 2013, ApJ, 778, 93
van den Bosch F. C., Aquino D., Yang X., Mo H. J., Pasquali A., McIntosh D. H., Weinmann S. M., Kang X., 2008, MNRAS, 387, 79
Vulcani B., Poggianti B. M., Finn R. A., Rudnick G., Desai V., Bamford S., 2010, ApJ, 710, L1
Wetzel A. R., Tinker J. L., Conroy C., 2012, MNRAS, 424, 232
Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013, MNRAS, 432, 336
Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, ApJ, 754, L29
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

This paper has been typeset from a TEX/LATEX file prepared by the author.