When should we treat galaxies as isolated?

Philip F. Hopkins,1⋆ Dušan Kereš,2 Chung-Pei Ma1 and Eliot Quataert1

1Department of Astronomy, University of California Berkeley, Berkeley, CA 94720, USA
2Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Accepted 2009 September 11. Received 2009 September 10; in original form 2009 February 12

ABSTRACT
Traditionally, secular evolution is defined as evolution of systems where the internal growth of structure and instabilities dominates the growth via external drivers (e.g. accretion and mergers). Most study has focused on ‘isolated’ galaxies, where seed asymmetries may represent realistic cosmological substructure, but subsequent evolution ignores galaxy growth and interactions. Large-scale modes in the disc then grow on a time-scale of the order of a disc rotation period (\(\sim 0.1–1\) Gyr). If, however, galaxies evolve cosmologically on a shorter time-scale, then it may not be appropriate to consider them ‘isolated’. We outline simple scalings to ask whether, under realistic conditions, the time-scale for secular evolution is shorter than the time-scale for cosmological accretion and mergers. We show that this is the case in a relatively narrow but important range of perturbation amplitudes corresponding to substructure or mode/bar fractional amplitudes \(\delta \sim 0.01–0.1\), the range of most interest for observed strong bars and most pseudo-bulges. At smaller amplitudes \(\delta \ll 0.1\), systems are not isolated: typical discs will grow by accretion at a comparable level over even a single dynamical time. At larger amplitudes \(\delta \gg 0.1\), the evolution is no longer secular; the direct gravitational evolution of the seed substructure swamps the internal disc response. We derive criteria for when discs can be well approximated as ‘isolated’ as a function of mass, redshift and disc stability. The relevant parameter space shrinks at higher mass, higher disc stability and higher \(z\) as accretion rates increase. The cosmological rate of galaxy evolution also defines a maximum bar/mode lifetime of practical interest, of \(\sim 0.1 t_{\text{Hubble}}(z)\). Longer lived modes will encounter cosmological effects and will decouple from their drivers (if they are driven).

Key words: galaxies: active – galaxies: evolution – galaxies: formation – galaxies: spiral – cosmology: theory.

1 INTRODUCTION
Isolated disc galaxies are prone to a number of important instabilities that play a major role in shaping observed late-type disc and bulge populations, with the most well known and well studied being the traditional bar and spiral instabilities. Both bars and spiral structure are ubiquitous in the local-disc population (Marinova & Jogee 2007; Menéndez-Delmestre et al. 2007; Barazza, Jogee & Marinova 2008), and their abundance appears comparable at higher redshifts (Sheth et al. 2003, 2008; Jogee et al. 2004). By amplifying small perturbations into coherent, long-lived, large-scale non-axisymmetric modes, these structures enable discs to evolve significantly – redistributing material in angular momentum and phase space – in a few orbital periods. As a consequence, observations and simulations indicate that these structures are important in shaping the cosmological evolution of disc sizes, scaleheights and the abundance, structural properties and mass fraction in ‘pseudo-bulges’ (disc-like bulges that result from angular momentum exchange in these modes), a population increasingly prominent in low-mass and later-type disc galaxies (e.g. Debattista et al. 2004; Kormendy & Kennicutt 2004; Weinzirl et al. 2009).

Traditionally, the growth and evolution of these global modes is referred to as ‘secular’ evolution: by definition, evolution that is slow relative to the local dynamical time. This contrasts with violent relaxation – seen in, for example, galaxy–galaxy major mergers – in which the potential fluctuates on short time-scales, and local instabilities, involving, for example, clumping, star formation and formation of bars on small scales (sub-kpc).

As a consequence, the secular evolutionary channel has, for the most part, been studied in the context of isolated galaxies. Given an isolated, self-gravitating stellar (or stellar+gas) disc that meets certain instability criteria, small non-zero amplitude in the large-scale modes that identify morphological bar and spiral patterns

⋆E-mail: phopkins@astro.berkeley.edu
(characteristic wavelength of the order of the disc length) will grow exponentially on a time-scale of a few orbital periods (see e.g. the discussion in Binney & Tremaine 1987). The evolution and dynamics of these modes have been well studied in idealized cases of isolated discs with properties similar to the Milky Way (MW), but by design bar or spiral wave unstable (Schwarz 1981; Athanassoula et al. 1983; Pfenniger 1984; Weinberg 1985; Combes et al. 1990; Hernquist & Weinberg 1992a; Friedli & Benz 1993; Patsis & Athanassoula 2000; Athanassoula 2002a,b; Athanassoula & Misiriotis 2002; Berentzen et al. 2003; 2004; Mayer & Wadsley 2004; Kaufmann et al. 2007; Weinberg & Katz 2007b; Foyle, Courteau & Thacker 2008). This work has informed subsequent studies of the role of secular evolution in shaping galaxy sizes, dynamics and morphology.

However, in a cold dark matter (ΛCDM) cosmologies, structure grows via continuous accretion and mergers. Although major mergers are rare, both theoretical calculations and observations suggest that minor mergers are ubiquitous, and accretion of new cold gas is rapid in low-mass galaxies (Maller et al. 2006; Woods, Geller & Barton 2006; Barton et al. 2007; Woods & Geller 2007; Stewart et al. 2008). Together with the typical substructure present in ΛCDM haloes (Gao et al. 2004; Taylor & Babul 2004), this suggests the concern that there may not be in practice such a thing as an ‘isolated’ galaxy at the level of interest.

More recent studies of secular evolution have therefore focused on more realistic scenarios, exploiting merger histories from cosmological simulations in semi-idealized studies of single galaxies (Bournaud & Combes 2002; Benson et al. 2004; Berentzen & Shlosman 2006, 2008; Gauthier, Dubinski & Widrow 2006; Curir, Mazzei & Murante 2007; Kaufmann et al. 2007; Kazantzidis et al. 2008; Romano-Diaz et al. 2008). These simulations again reveal bars and spiral structure to be prominent – arguably more so than in isolated simulations – but it is less clear whether their formation and evolution can be attributed to the same secular processes at work in isolated systems, or whether they are driven systems owing to substructure and accretion in the galaxy disc and halo.

The important question for models is that can any galaxy in a realistic cosmological context still be approximated as ‘isolated’ for certain purposes? If so, in what regimes as a function of redshift, galaxy mass and internal properties is this applicable? What are the corresponding implications for the interpretation of bar fractions and lifetimes? Ultimately, what does this imply for the importance of isolated secular evolution in driving the evolution of galaxies and formation of bulges?

In this paper, we attempt to address these questions by means of a simple comparison of cosmological accretion rates and characteristic time-scales for secular evolution. This approach allows us to identify the regimes where galaxies can be safely considered ‘isolated’ versus where cosmological effects may not be negligible. We show that there is an interesting regime of secular modes with fractional mass/amplitude of ∼0.1 where the secular growth mode dominates and the isolated galaxy approximation is good (Section 2). We show how this scales with galaxy mass, redshift and disc-stability properties (Section 3), and identify some basic consequences for the lifetimes of large-scale modes in discs (Section 4). Our goal is not a definitive description of secular evolution, but rather to provide a set of simple initial constraints to provide context for more detailed studies of the interesting parameter space.

Throughout, we adopt an ΩΛ = 0.3, ΩM = 0.7, h = 0.7 cosmology, but our conclusions are not sensitive to the choice within the range allowed by present observations (e.g. Komatsu et al. 2009).

2 SECULAR EVOLUTION VERSUS COSMOLOGICAL EVOLUTION

Consider an ‘initial’ equilibrium, axisymmetric disc-halo system at time t = 0. In this limit, the system will not evolve any non-axisymmetric modes. Therefore, introduce a non-axisymmetric perturbation to the disc potential of amplitude

$$\delta_0 = \frac{\delta \phi}{\phi}.$$  (1)

We are specifically interested in global models, so $\phi \sim GM/R$ is the potential of interest (where $M$ is the disc+enclosed halo mass and $R$ is a characteristic effective radius/size length). The precise meaning of the perturbation $\delta \phi$ differs depending on the mode(s) of interest and configuration. For example, in idealized N-body simulations, this typically corresponds to shot noise. However, in realistic cosmological settings this will correspond to substructure in the disc or halo, with $\delta \phi \sim GM/R$ (where $m$ is the substructure mass and $r$ its ‘initial’ distance). The relevant numerical prefactor will depend on the orbit, phase-space structure and mode (e.g., for a bar, the desired quantity is the time-averaged contribution to the $m = 2$ mode at radius $\sim R$ in the corotating frame); for our purposes, the scaling (not absolute value) of $\delta$ is most important.

At early times (before saturation), this non-axisymmetric term will be amplified internally and grow roughly exponentially:

$$\delta(t) = \delta_0 \exp(t/t_0),$$  (2)

where $t_0$ is the effective secular time-scale, which is typical of the order of a few orbital times (again, this is for global modes, not local; see e.g. Holley-Bockelmann, Weinberg & Katz 2005; Weinberg & Katz 2007a,b). This growth time has been the focus of a considerable number study, and is one of the many important results of isolated disc studies. For example, for a disc bar in a strongly unstable bulge-free MW-like disc, Dubinski, Berentzen & Shlosman (2009) show that equation (2) is a good approximation to the behaviour in simulations, with $t_0 = 8\pi/\kappa \approx 2.83 P_d^4$ (where $\kappa$ is the epicyclic frequency $= 2\pi^2\pi P_d^{-1}$ for a constant circular velocity disc and $P_d = 2\pi R/V_c$ is the disc circular period at its effective radius). Klypin et al. (2008) find a similar $t_0 \approx 3–5 P_d$ for thin, bulge-free MW-like discs (albeit with a much larger $t_0 \sim 10–30 P_d$ for thick $H/R \gtrsim 0.5$ discs; see also Colín, Valenzuela & Klypin 2006). Martinez-Valpuesta, Shlosman & Heller (2006) see time-scales from ∼2.5 to 10 $P_d$, depending on whether the bar growing is an initial mode or a secondary (post-buckling) mode. A similar range of time-scales is found (with considerable galaxy-to-galaxy variation) in live cosmological haloes in Berentzen & Shlosman (2006).

For less cosmologically motivated but more general and analytically tractable disc mass-profiles, Athanassoula & Sellwood (1986) find typical $t_0 \sim 1.0–6.7 P_d$ for realistic halo mass fractions ∼1/4–1/2 (fraction of the total mass owing to the halo at < R) and scale-heights $H/R \sim 0.1$. Narayan, Goldreich & Goodman (1987) and Shu et al. (1990) obtain $t_0 \sim 0.8–1 P_d$ for gas discs with an outer Lindblad resonance at $R \gtrsim R_0$ (of interest for global modes here) and no halo.

More stable systems will evolve more slowly; for the sake of generality we define

$$t_0 = N_{\text{disc, Periods}} \times P_d \equiv \frac{1}{1 - \chi_{\text{eff}}} P_d,$$  (3)

where $\chi_{\text{eff}}$ is an effective stability parameter: $\chi_{\text{eff}} \sim 0$ represents typical, cosmologically realistic discs maximally unstable to large-scale modes, which will evolve on a single orbital time; and $\chi_{\text{eff}} > 1$
systems are stable and experience only oscillations, rather than amplifying modes.\textsuperscript{1} Note that, formally speaking, $\chi_{\text{eff}} < 0$ is allowed. For certain bar configurations, for example, $t_0 \sim 0.7 P_d$ has been obtained (see e.g. Adams, Ruden & Shu 1989; Earn & Sellwood 1995), or even, for spiral structure in the weak winding approximation, $t_0 \sim 0.4 P_d$ (Toomre 1981). However, those situations all involve no halo and an infinitely thin disc and somewhat different matter profiles from what are observed in typical discs. For moderate halo contributions or disc thickness, $t_0$ is unlikely to be smaller than $P_d$ by any but a small factor ($t_0 \sim 0.7-0.8 P_d$), a small difference relative to the uncertainties in other quantities calculated here. The MW-like examples above illustrate that $\chi_{\text{eff}} \approx 0.5-0.75$ is probably the case of greatest interest for realistic disc plus halo systems, even for strongly unstable systems. To be conservative, however, we will adopt $\chi_{\text{eff}} = 0$ for all numerical estimates, unless explicitly otherwise specified.

However, galaxies are not static, and two things will happen that might compete with this internal self-amplification. (1) The substructure itself can dynamically evolve, driving stronger perturbations and/or merging. (2) New mass of magnitude comparable to the disc mode can be accreted/merged. If either of these occurs on a time-scale shorter than $t_0$ (the effective secular time-scale), the system should not be considered ‘isolated’ for purposes of secular evolution.

Consider Case (1), the dynamical evolution of the substructure itself. Given some substructure/perturbation of mass fraction $\delta$ at some initial radius of interest $r$, the orbit will decay on a time-scale of the order of the dynamical friction time; correspondingly, the perturbation $\delta \propto \delta \phi \propto r^{-1}$ will grow on the same time-scale.\textsuperscript{2} Strictly speaking, dynamical friction does not dominate angular momentum loss at small radii; rather, resonant tidal interactions act more efficiently (Barnes & Hernquist 1992). However, properly calibrated, the dynamical friction time is not a bad approximation (e.g. Boylan-Kolchin, Ma & Quataert 2008). For an isothermal sphere or Mestel (1963) (flat rotation curve) disc, this time is simply

$$t_{\text{eff}} = \frac{R/V_c}{2 \beta \ln \Lambda} \frac{M_{\text{acc}}(r)}{m} \frac{R}{R} \approx \frac{0.2}{\delta_0 \ln \Lambda} P_d$$

(5)

where the equality on the right comes from the definitions of $\delta_0$ and $P_d$.\textsuperscript{3} Since we are considering the magnitude of the perturbation relative to the disc, the time here scales with the disc dynamical time at fixed $\delta_0$ (as opposed to e.g. the Hubble time for halo–halo orbital decay at large radii).

The left-hand panel of Fig. 1 compares this time-scale to the secular evolution time-scale $t_0$. For representative purposes, we assume a ‘maximally unstable’ $t_0 = P_d$ ($\chi_{\text{eff}} = 0$) MW-like disc scale of the order of the dynamical friction time; correspondingly, the perturbation $\delta \propto \delta \phi \propto r^{-1}$ will grow on the same time-scale.\textsuperscript{2} Strictly speaking, dynamical friction does not dominate angular momentum loss at small radii; rather, resonant tidal interactions act more efficiently (Barnes & Hernquist 1992). However, properly calibrated, the dynamical friction time is not a bad approximation (e.g. Boylan-Kolchin, Ma & Quataert 2008). For an isothermal sphere or Mestel (1963) (flat rotation curve) disc, this time is simply

$$t_{\text{eff}} = \frac{R/V_c}{2 \beta \ln \Lambda} \frac{M_{\text{acc}}(r)}{m} \frac{R}{R} \approx \frac{0.2}{\delta_0 \ln \Lambda} P_d$$

(5)

where the equality on the right comes from the definitions of $\delta_0$ and $P_d$.\textsuperscript{3} Since we are considering the magnitude of the perturbation relative to the disc, the time here scales with the disc dynamical time at fixed $\delta_0$ (as opposed to e.g. the Hubble time for halo–halo orbital decay at large radii).

The left-hand panel of Fig. 1 compares this time-scale to the secular evolution time-scale $t_0$. For representative purposes, we assume a ‘maximally unstable’ $t_0 = P_d$ ($\chi_{\text{eff}} = 0$) MW-like disc

\textsuperscript{1} Under certain restrictive circumstances, our $\chi_{\text{eff}}$ here is analogous to the Toomre Q or X parameter $X \equiv \kappa^2 R / (2 \pi m G \Sigma)$ or a (renormalized) Ostriker–Peebles criterion (proportional to the ratio of rotational kinetic to potential energy). For example the bar in a two-dimensional Kuz’m in disc approximation presented in Athanassoula & Sellwood (1986), we can translate their equation (3) to obtain

$$\chi_{\text{eff}} = 0.3 + 1.1 \left( \frac{f_{\text{halo}} + f_{\text{buige}}}{1/3} - 1 \right) + 0.6 \left( \sqrt{\frac{H}{0.1 R}} - 1 \right)$$

(4)

in physical terms of the disc thickness $H/R$ and halo plus bulge (non-disc) mass fraction inside $R$. The definition in equation (3) is not, however, meant to represent specific instabilities but to allow for general large-scale disc modes with a characteristic growth time/stability criterion.

\textsuperscript{2} Strictly speaking, realistic cosmological perturbations grow continuously, so an ‘initial’ radius is ambiguous. However, there is still some $\delta \phi$ that scales as described at a given instantaneous $r$, and this is what ultimately enters into the equations derived. Also, in practice, such modes – where induced by substructure – often appear suddenly (i.e. in a time $< P_d$ when $r \sim R$; this is because at larger radii, the net non-axisymmetric $\delta \phi$ contribution is suppressed by a Poisson $\sim N^{-1/2} \sim R^{-3/2}$ term. In simulations, for example, perturbations are typically dominated by a few close passages of clumps/substructure where $r \sim R$ (although these may be from longer radial orbits; see Velazquez & White 1999; Bournaud & Combes 2002; Gauthier et al. 2006; Hopkins et al. 2008b; Kazantzidis et al. 2008). In any case, since our derivations rely on $\delta$, rather than $r$ explicitly, this is not a large source of uncertainty.

\textsuperscript{3} In detail, $\beta$ is a constant that weakly depends on the mass profile and velocity isotropy: $\approx 0.428$ for an isotropic isothermal sphere and $\approx 0.32$ for a thin Mestel (1963) disc averaged over random inclinations (used in equation 5). The Coulomb logarithm is approximately $\Lambda = 1 + 1/\delta_0$ (Boylan-Kolchin et al. 2008; Jiang et al. 2008). For Fig. 1, we use the fitting functions from Boylan-Kolchin et al. (2008), with appropriate eccentricity and orbital parameter dependence, rather than the simplified equation (5), but the results are similar on average.
with $P_A = 2\pi \times 5\, \text{kpc}/200\, \text{km s}^{-1} \approx 160\, \text{Myr}$, and total stellar mass $= 5 \times 10^{10} M_\odot$. This is easily generalized; $P_A$ (at the scale $\sim R$ of the disc itself) appears to be independent of mass in observed discs (e.g. Courteau et al. 2007, and references therein). We plot the results assuming such a disc exists at redshift $z = 1$, but the qualitative scalings are similar at redshift $z = 0$, and we will show the redshift dependence explicitly below. We compare the dynamical friction time $t_{df}$; here we show the results using the full orbital parameter-dependent fits from simulations in Boylan-Kolchin et al. (2008), which allows us to quote the $\pm 1\sigma$ range of $t_{df}$ from the range of orbits observed in cosmological simulations (Benson 2005; Khochfar & Burkert 2006). Using the simpler formula in equation (5) is similar to the median expected.

Comparison of Fig. 1 or equations (3) and (5) shows that the dynamical evolution of the perturbation is more rapid than the internal response for mass ratios larger than

$$\delta_{\text{crit, at}} = \frac{1 - \chi_{\text{eff}}}{4 \pi \beta \ln \Lambda} \sim 0.2 (1 - \chi_{\text{eff}}).$$

This is ultimately an obvious regime; when $\delta \phi / \phi \sim 1$, direct evolution dominates the potential fluctuations. We denote this as the ‘major merger regime’: in the case where $\delta$ corresponds to some substructure, this clearly requires a mass ratio $\mu \gtrsim 0.2$ with $r \sim R_{200}$, i.e. close passages of major companions. Note though that this does not have to be a merger. For example, a sufficiently strong disc-fragmentation event will be similar. Physically, this is still dynamically distinct from secular evolution (from e.g. bars, etc.) – it will ‘look like’ a merger inside the disc (see e.g. Elmegreen, Bournaud & Elmegreen 2008).

Now consider Case (2): new growth/perturbations/mergers. Note that we are no longer considering the evolution of individual perturbations, but the time between new perturbations of the same or greater magnitude. If this is $\ll t_0$, then the system is not isolated. A lower limit to this is given by the rate of baryonic accretion/merging on to the disc (if accreted systems retain some dark matter, they will represent larger perturbations, but there is at least a lower limit in the mass added in baryons to explain the disc mass). Detailed analyses of these rates have been discussed extensively in the literature (see e.g. Brown et al. 2007; Genel et al. 2008; Guo & White 2008; Stewart et al. 2008; Wetzel, Cohn & White 2009). Here, we use a simple semi-empirical model to define some of the relevant scalings; for more discussion, see Hopkins et al. (2009a). A variant of the model, based on subhalo–subhalo merger rates, is also described in detail in Hopkins et al. (2008a). Following Stewart et al. (2009b), we begin with dark matter halo merger trees (here from Fakhouri & Ma 2008). Empirical halo occupation models and other observations constrain the average galaxy mass per host halo (or subhalo) mass, with little scatter – so at a given instant we simply populate the haloes with galaxies. Specifically, we assign stellar mass given the fitted $M_\star (M_{\text{halo}} | z)$ from Conroy & Wechsler (2009) and gas mass given the fits to $M_\text{gas} (M_\star | z)$ from Stewart et al. (2009a) for the observations used in the fits, see references therein and Bell & de Jong 2001; Erb et al. 2006; Fontana et al. 2006; Pérez-González et al. 2008).

The uncertainties in this modelling methodology will be discussed in detail in Hopkins et al. (2009b), but for our purposes they are relatively small (a factor of $\sim 2$ uncertainty in the merger rate near $\sim \chi_{\text{int}}$, owing to a combination of uncertainty in $M_\star (M_{\text{halo}})$ and the halo–halo merger rate) at $z < 2$, because it is primarily the shape of the galaxy–halo mass correlation (rather than e.g. its absolute normalization) that affects galaxy–galaxy merger rates.\footnote{The merger rates from this model as used here can be also obtained as a function of, for example, galaxy mass, mass ratio and redshift from the ‘merger-rate calculator’ script publicly available at http://www.cfa.harvard.edu/~hopkins/Site/mergercalc.html.}

Note, however, that the uncertainties grow rapidly at higher redshifts, owing to the lack of empirical constraints. Evolving the system forward some small increment in time, we can ‘add up’ the mergers (in detail, we add a dynamical friction ‘delay’ time between each halo–halo merger and subsequent galaxy–galaxy merger, with the formulae from Boylan-Kolchin et al. 2008). This gives merger rates; but also, knowing the new halo mass (after accretion/growth in this time interval), the empirical halo occupation constraints define the ‘expected’ galaxy mass for the updated halo mass. We simply assign whatever galaxy mass growth is needed to match this (not already brought in by mergers) to ‘accretion’. Note that this is a lower limit to the accretion rate, reflecting net accretion (outflows may remove mass, requiring more new gas inflow).

The middle panel of Fig. 1 shows the relevant time-scale for both mergers (median time $\Delta t$ between mergers with baryonic mass ratio $\mu \equiv M_{\text{bar}, z}/M_{\text{bar}, 1} > \delta_0$) and accretion ($\Delta t$ for the disc to grow via accretion by a mass fraction $> \delta_0$). Accretion tends to be the dominant growth channel (relative to e.g. minor mergers), for all but the most massive galaxies (where gas accretion is ‘quenched’). As a result, the time between new mergers may be long, but at sufficiently low $\delta_0$, growth by accretion is more rapid than internal disc response. We denote this by the ‘accretion regime’. Again, the behaviour is easily understood: if one is interested in evolution at the $\ll 10$ per cent level, then galaxies cannot be considered isolated for even a single dynamical time, as they will grow by more than this amount in that time.

The relevant criterion can be roughly estimated as follows: to very crude approximation, fractional galaxy growth rates scale as $\sim \alpha / t_{\text{accretion}}$, where $\alpha$ is weakly redshift-dependent but non-trivially mass-dependent with $\alpha \sim 0.2$ for a MW mass halo at $z = 1$ (i.e. an assumed galaxy mass of $5 \times 10^{10} M_\odot$). For such a system, as pictured in Fig. 1, the galaxy will grow by a fraction $> \delta_0$ in the time $t_0$ (secular response time) for perturbation amplitudes below

$$\delta_{\text{crit, acc}} = \frac{\alpha}{1 - \chi_{\text{eff}}} \times \frac{P_A}{t_{\text{df}}(z)} \sim 0.003 (z = 0).$$

Fig. 1 considers the ‘maximally unstable’ ($\chi_{\text{eff}} = 0$) case, such that $t_0 = P_A$. If the stability parameter is higher (larger $t_0$), the regime of effective ‘isolation’ will be more restricted. Fig. 2 illustrates the parameter space as a function of the effective disc-stability parameter $\chi_{\text{eff}}$ (recall, this is simply defined relative to the number or orbits needed to grow the mode of interest). Above some critical $\chi_{\text{eff}}$ (here $\chi_{\text{eff}} \sim 0.75$, i.e. $N_{\text{orbits}} = 4$ or $t_0 \gtrsim 0.5$ Gyr for a MW-like disc), the secular time-scale is always longer than the other time-scales above. This is simply the statement that discs are not ‘isolated’ for time-scales $\gtrsim$ Gyr, especially at high redshift.

3 DEPENDENCE ON GALAXY MASS AND REDSHIFT

Fig. 3 shows how the regime of secular evolution depends on galaxy mass and redshift. First, we consider the same comparison at $z = 0$ as a function of galaxy mass. Observations indicate that $P_A$ is nearly mass-independent at the disc-effective radii of interest for global models (Bell & de Jong 2001; Shen et al. 2003; Courteau et al. 2007). Given equation (5), the same is true for dynamical evolution.
of individual perturbations (at fixed $\delta_0$). However, accretion and merger rates scale significantly with mass. At low masses, merger rates are low, but accretion rates are high. At high masses, accretion rates drop rapidly (consistent with zero at $M_{\text{gal}} \gg 10^{11} \text{ M}_\odot$), but merger rates increase, leaving almost no range of perturbation in which secular processes are relevant (right-hand column of Fig. 3). Both effects are seen in a variety of models and observations (Maller et al. 2006; Noeske et al. 2007; Guo & White 2008; Kitzbichler & White 2008; Bundy et al. 2009; Kereš et al. 2009; Parry, Eke & Frenk 2009; Stewart et al. 2009b). The mass dependence is important even over a relatively narrow mass range – for example, note that our previously assumed MW-like mass of $5 \times 10^{10} \text{ M}_\odot$ (Fig. 1), being a factor of $\sim 2$ smaller than the $10^{11} \text{ M}_\odot$ case shown here, has correspondingly more rapid accretion rates (between the $10^{10} \text{ M}_\odot$ curve and $10^{11} \text{ M}_\odot$ curve).

Again, we emphasize that we are using baryonic mass ratio $\mu$ here – this is a minimum, as it reflects the most densely bound material that will survive to perturb the galaxy (an individual merger may ‘begin’ at larger $\delta$ including dark matter, or smaller $\delta$ at large radii, but orbital decay and stripping will tend to saturate it at $\delta \phi/\phi \sim \mu$, with a rate of such new events from mergers as shown; see e.g. Kazantzidis et al. 2008). Low-mass galaxies are observed to be more dark matter dominated, so if this can be conserved, the relevant rates will not decrease as rapidly with stellar mass; however, modelling this requires more detailed knowledge of cosmological orbits, stripping and internal galaxy structure.

For each mass, Fig. 3 shows how the regime of secular evolution depends on redshift. To the lowest order, accretion timescales evolve with the Hubble time (fitting directly, accretion rates $\propto (1 + z)^2$; see Stewart et al. 2009b). Observations of the baryonic Tully–Fisher and size mass relation suggest that $P_\delta$ (or equivalently at fixed mass, disc sizes) evolves weakly from $z = 0$–2 (Flores et al. 2006; Trujillo et al. 2006; Kassin et al. 2007; Toft et al. 2007; Akiyama et al. 2008; Somerville et al. 2008). Moreover, theoretical models that include the well-established dependence of halo concentration on redshift (see e.g. Bullock et al. 2001; Wechsler et al. 2002) predict a similar weak scaling (Somerville et al. 2008). Parameterizing as $P_\delta \propto (1 + z)^{-\beta_\delta}$, these observations constrain $\beta_\delta = 0.0$–0.6. In Fig. 3, we conservatively adopt $\beta_\delta = 0$ (i.e. $P_\delta$ independent of redshift), but we show how the results would change if we allowed the maximum observationally inferred evolution, $\beta_\delta = 0.6$. It makes a small difference, but does cancel some of the redshift evolution in the relevant parameter space. Even in the extreme case of a simple $P_\delta \propto t_{\text{Hubble}}$ scaling (Mo, Mao & White 1998), some,

Figure 2. Parameter space of regimes in Fig. 1 for the same $\sim L_*$ system as a function of perturbation amplitude and effective disc stability (speed of growth of secular modes of interest).

Figure 3. Same as the right-hand panel of Fig. 1 ($x_{\text{eff}} = 0$), as a function of galaxy mass and redshift. In low-mass ($\lesssim 10^{10} \text{ M}_\odot$; Left-hand panel) galaxies, merger rates are low, but accretion rates are rapid – secular responses at the $<10$ per cent level compete with cosmological disc growth. At intermediate masses $\sim L_*$($\sim 10^{11} \text{ M}_\odot$; Middle panel) accretion and minor mergers occur with comparable rates. At high masses ($\gtrsim 10^{11} \text{ M}_\odot$; right-hand panel) accretion rates are low (cooling is inefficient) but merger rates grow rapidly – secular responses at the $<10$ per cent level compete with mergers. As a function of redshift, disc dynamical times scale weakly, but merger and accretion rates increase, leaving less of parameter space in which discs can be considered ‘isolated’ for the internal response time. Error bars mark the range between the internal response time if disc sizes do not evolve with redshift ($\beta_\delta = 0$; dashed lines) and if they evolve at the maximum rate constrained by observations ($\beta_\delta = 0.6$; lower bar).
but not all of the evolution is negated (at $z < 2$, merger and cooling rates evolve as $\propto (1 + z)^2$, $1/a_{\text{Hubble}}$ as $\propto (1 + z)$).

In Fig. 3, the critical amplitude below which the ‘accretion regime’ pertains scales roughly as $\delta_{\text{crit, acc}} \propto (1 + z)^{1.5–2.0}$, while $\delta_{\text{crit, acc}} \sim$ constant. This is an approximation over the entire range $z = 0–2$; in fact at the lowest redshifts ($z \lesssim 0.2$), the falloff in $\delta_{\text{crit, acc}}$ is somewhat more rapid (as e.g. the Universe’s acceleration term becomes important). As a consequence, the range of $\delta_0$ over which ‘isolation’ is a good approximation decreases with the increasing redshift.

Fig. 4 summarizes the parameter space as a function of galaxy mass and stability parameter $\chi_{\text{eff}}$, at $z = 0$ and 1 and $z = z_{\text{form}}(M_{\text{gal}})$. We define $z_{\text{form}}(M_{\text{gal}})$, the galaxy assembly time, as the redshift when each galaxy reaches half its $z = 0$ mass, according to our simple growth model. To the extent that secular modes are considered important in this formation process, this is an interesting time-scale.

Simulations find that star-forming galaxies accrete most of their mass along a couple of dynamically coherent, clumpy filaments; as such they are dynamically important for large-scale disc modes (Kereš et al. 2005, 2009; Dekel & Birnboim 2006; Dekel et al. 2009).

If, however, accretion were perfectly smooth, axisymmetric and restricted to large radii (without migration of new material inwards), then it might be valid to ignore it in studying secular modes even when accretion rates are large. To represent this possibility, Fig. 5 recalculates Fig. 4, but ignores accretion. At low masses, merger rates are sufficiently low that the isolated regime extends to smaller mass ratios $\delta < 0.01$.

### 4 IMPLICATIONS FOR MODE ‘LIFETIMES’

The cosmological evolution of galaxies also has important implications for mode ‘lifetimes’. Since $t_0 = 1/(1 - \chi_{\text{eff}}) P_{\text{d}}$, there is clearly some $\chi_{\text{eff}}$ at each redshift above which $t_0$ is larger than any of the competing time-scales for all $\delta_0$. Modes with larger $\chi_{\text{eff}}$ are still formally unstable, but the time/number of orbits to amplify the mode becomes sufficiently long that these modes should be considered cosmologically dynamical objects. Fig. 6 shows this maximum $\chi_{\text{eff}}$ as a function of redshift (for $\sim 10^{11} M_\odot$ galaxies where this is maximized, as seen in Fig. 4). At $z \gtrsim 1$, this corresponds to modes that become long-lived above $\delta_0 = 0.05$ but rapidly cool at $\delta_0 \gtrsim 0.75$.
growing in \( \lesssim \) a couple \( P_d \); at \( z \gtrsim 2 \), however, even \( \chi_{\text{eff}} = \ll 1 \) systems (those where modes grow on a time-scale \( \sim P_d \)) can be in the ‘accretion regime’, as discussed above. Recall, simulations suggest that even cold, bulge-free MW-like discs have effective \( \chi_{\text{eff}} \sim 0.5–0.75 \) (Dubinski et al. 2009, and references therein). This high-\( \chi \) behaviour is directly related to observations showing that disc orbital periods at high redshifts become comparable to the Hubble time (see e.g. Flores et al. 2006; Kassin et al. 2007; Toft et al. 2007; Shapiro et al. 2008; van Starkenburg et al. 2008).

At \( \chi_{\text{eff}} \) less than the values above, secular modes can grow ‘in isolation’ from some \( \delta_{\text{in}} \). Typically, these will grow rapidly and saturate at some \( \delta_{\text{f}} \approx 1 \). However, if an isolated mode then survives stably at an amplitude \( \delta_f \) for a lifetime much longer than the other time-scales compared here, then various cosmological effects may have important consequences. For example, if a disc bar saturates and survives with some \( \delta_f \gtrsim 0.4 \) (Dubinski et al. 2009), in some number of dynamical times the galaxy will grow by this much. Essentially, cosmological growth may ‘catch up’ to the saturated mode and could affect it. Of course, the mode could continue growing with the galaxy or be robust to these effects; our point is that continuing to treat such a mode in isolation may not necessarily be a good approximation over much longer time-scales. Moreover, if stable modes can survive for a time-scale much longer than, for example, the relevant dynamical friction times at \( \delta_{\text{f}} \), then the presence of those modes mode (the duty cycle) will decouple from that of their drivers (if they were initially driven). In, for example, the case of minor mergers, this is the statement that new mergers and/or the destruction of the original driving satellites will wipe out the ‘memory’ of the drivers, while the bar survives.

Taking the minimum of the non-secular time-scales of interest (e.g. accretion and merger time-scales in Fig. 3), at whatever amplitude \( \delta \) maximizes this time-scale, gives the maximum relevant ‘isolated’ mode lifetime. This is clearly a function of mass; we consider here the \( \sim 10^{11} \text{M}_\odot \gtrsim (\sim L_\odot) \) case of greatest interest, both as a MW-like system and because Fig. 4 demonstrates that this is where such a time-scale (the ‘isolated’ regime) is maximized. Fig. 7 plots this time-scale versus redshift. We show this both for the assumption that \( P_d \) does not evolve (\( \beta_d = 0 \)) and the maximum observationally constrained evolution (\( \beta_d = 0.6 \)). We compare a constant fraction (\( \sim 0.1 \)) of the Hubble time – this appears to be a good approximation, on average (there will of course be scatter galaxy-to-galaxy in accretion and merger rates, leading to typical factor \( \sim 2 \) scatter in the relevant time-scale here).

5 DISCUSSION

Under typical cosmological conditions, global ‘secular’ evolution – narrowly defined as evolution by internal amplification of large-scale disc modes in effectively isolated galaxies – only occurs in a restricted range of parameter space (Figs 1 and 2). If the perturbation mode of interest has a fractional amplitude \( \lesssim 0.1 \), what we call the ‘accretion regime’, then the disc will grow by accretion by a comparable amount in even a single dynamical time; the isolated approximation is clearly not valid. This threshold is around an amplitude \( \delta_{\text{acc},\text{iso}} \gtrsim 0.002(1 - \chi_{\text{eff}})^{-1}(1 + z)^{1.5} \sim 10^{11} \text{M}_\odot \) galaxies (slightly lower at \( z < 0.2 \) or \( \delta_{\text{acc},\text{iso}} \gtrsim 0.005(1 - \chi_{\text{eff}})^{-1}(1 + z)^{1.5} \sim 10^{10} \text{M}_\odot \) systems. At the opposite extreme, seed ‘perturbations’ of fractional amplitude \( \delta_0 > \delta_{\text{acc},\text{iso}} \gtrsim 0.2 \) lead to non-secular evolution – the perturbations’ own gravitational evolution will dominate the internal response (this is obvious in the case of e.g. galaxy–galaxy major mergers or massive disc-fragmentation events, where the evolution of the merger/clumps drives violent relaxation).

The relevant parameter space depends on galaxy mass (Figs 3–5). Although halo growth is nearly mass-independent (Fakhouri & Ma 2008; Guo & White 2008; Stewart et al. 2009b), galaxy-growth histories are not [the function \( M_{\text{halo}}(M_{\text{gal}}) \) is non-trivial]. At high masses (\( M_{\text{gal}} \gtrsim 10^{11} \text{M}_\odot \)) galaxy–galaxy merger rates are high such that systems are rarely ‘isolated’ over the time-scales of interest for secular evolution. However, there are few disc at these masses, so secular evolution is not expected to be a dominant process. At low masses (\( \sim 10^{10} \text{M}_\odot \)) merger rates are low (in terms of galaxy–galaxy baryonic mass ratios; including dark matter, they may remain high) but accretion rates are high; systems can be effectively approximated as isolated for only a couple of orbits in the regime of amplitudes \( \sim 0.03–0.2 \). Moreover, although such galaxies are mostly disc (\( B/T \ll 1 \)), they are increasingly dark matter dominated which helps stabilize them to the development of secular modes (see e.g. Persic, Salucci & Stel 1996; Mihos, McGaugh & de Blok 1997; Bell & de Jong 2001; Borriello & Salucci 2001). Galaxies may be ‘most isolated’, and so traditional secular evolution most relevant, between these regimes, i.e. in galaxies somewhat below \( \sim L_\odot \). That this occurs at masses only somewhat below where mergers become efficient is also interesting; there may be a relatively rapid regime (as galaxies approach and cross \( \sim L_\odot \) in mass) in which today’s galaxies transition from accretion-dominated, secularly stable (dark matter dominated) discs to secularly unstable (self-gravitating) discs, which could quickly amplify \( \sim 10 \) per cent amplitude perturbations into very strong bars and build significant pseudo-bulges, until later mergers destroy the remains of the disc and build massive classical bulges.

This has important implications for the lifetimes of secular processes of interest. The above comparisons assume discs where the internal response occurs over a single orbital period; if the systems have higher effective stability (i.e. secular responses build more slowly), then the regime where they can be considered isolated for this time shrinks. Large-scale modes that require more than a few disc periods to self-amplify at low redshift, or more than just a single

**Figure 7.** The maximum ‘isolated’ lifetime of large-scale modes (e.g. disc bars). This is the largest time-scale (marginalizing over \( \delta \) and \( M_{\text{gal}} \)) shorter than the competing (non-secular) time-scales in, for example, Fig. 3. Variations within populations contribute factor \( \sim 2 \) scatter from object-to-object. We show a fixed fraction of the Hubble time for comparison. Evolution on longer time-scales will compete with cosmological effects. The duty cycle will decouple from driving: even if all such modes were driven by, for example, minor mergers, there will be no correlation between the presence of the modes and the presence of companions.
disc period at high redshift ($z > 2$), should be considered cosmologically dynamical systems (Fig. 6) – the galaxy grows comparably over this self-amplification time-scale. Indeed, various observations of disc sizes and structure suggest that discs are sufficiently thick or have sufficient bulge fractions such that internal response times are in this interesting range (Barteldeeves & Dettmar 1994; de Grijz, Peletier & van der Kruit 1997; Bell & de Jong 2001; Gilmore, Wyse & Norris 2002; McGaugh 2005; Wyse et al. 2006; Courteau et al. 2007).

Even if modes can evolve/self-amplify quickly such that a bar will grow efficiently and saturate at some final amplitude, these competing time-scales define a maximum ‘isolated’ lifetime for that saturated mode that is of interest, $\sim 0.1 Hubble$ (Fig. 7). There has been substantial debate regarding the lifetime of stellar bars in discs; but if modes live stably in isolation for longer than this time, they will encounter significant cosmological effects including, for example, significant new disc growth and mergers. Indeed, most studies do agree that lifetimes in isolation are at least this long (see e.g. Weinberg 1985; Hernquist & Weinberg 1992b; Friedli, Benz & Kennicutt 1994; Athanassoula 2002a; Kaufmann et al. 2007). Evolution of modes on longer time-scales (e.g. some self damping or buckling processes) should ideally be considered in a live cosmological context – the time in isolation may strengthen modes against external effects, but various studies have found that a moderate level of new gas accretion or passages of new substructure can dramatically change mode evolution, both exciting and destroying bars and spiral waves (see Bournaud & Combes 2002; Berentzen et al. 2003, 2004, 2007; Athanassoula, Lambert & Dehnen 2005; Foyle et al. 2008); not to mention that the presence of pre-existing strong bars may in turn affect these accretion/merger processes.

Moreover, if modes live this long, their duty cycles will decouple from those of their drivers. Even if, for example, all large-scale bars were initially driven by encounters with satellite galaxies (minor interactions) and if the isolated lifetime were much longer than this value, there would be no surviving correlation between the presence of bars and such companions. There has been considerable observational debate regarding whether or not strongly barred galaxies exhibit any strong preference for minor companions; certainly there are at least many such galaxies without close neighbours (see Elmegreen, Elmegreen & Bellin 1990; Odewahn 1994; Moles, Marquez & Perez 1995; Marquez & Moles 1996; Li et al. 2009, and references therein). This may in fact be because strong bars are not driven; however, it could also be consistent with the hypothesis that all such bars were initially driven but are sufficiently long-lived. Constraints on bar lifetimes are needed to break the degeneracies.

The level of cosmological dynamics also has implications for the numerical considerations involved in simulations of ‘isolated’ systems. Properly following resonant self-interactions of bars may imply steep resolution requirements in N-body experiments (see e.g. Ceverino & Klypin 2007; Weinberg & Katz 2007a,b; Sellwood 2008). However, there are other properties for which increasing the resolution in idealized cases may not be a more accurate representation of reality. In terms of shot noise in the potential, for example, a model MW-like disc with $> 10^6$ particles will have potential fluctuations from smooth axisymmetry $\delta \phi / \phi < 1$ per cent over the spatial/time-scales of interest (disc size and dynamical time). In cosmological simulations, although the central regions of haloes are relatively smooth, even dark matter only simulations yield comparable or larger variations in the local potential/velocity dispersion at, for example, MW-like disc-effective radii (see Zemp et al. 2009). Even where smooth in space, such systems are not constant in time (as in idealized cases) at this level over several dynamical times. Moreover, inclusion of baryons (which are not stripped efficiently, unlike dark matter subhaloes which are efficiently destroyed at small radii and so do not ‘survive’ to contribute substructure inside the centres of haloes) enhances the clumpy, minor spatial substructure. In the MW, for example, the Large/Small Magellanic Cloud (LMC/SMC) system represents a real deviation from a smooth, axisymmetric potential at a level larger than this limit near the solar radius. Ideally, tracking the evolution of substructure at higher resolution should involve not just a larger number of particles, but cosmologically motivated descriptions of substructure and accretion.

Interestingly, at all redshifts, we find that traditional isolated ‘secular’ evolution is most applicable around perturbations of fractional amplitude $\sim 0.1$ per cent. This is a very interesting regime of parameter space: to the extent that it represents a fractional amplitude of substructure/accretion flows, it is a channel by which haloes and low-mass galaxies gain much of their mass (e.g. Governato et al. 2007; Kazantzidis et al. 2008; Stewart et al. 2008). Moreover, ‘pseudo-bulges’ associated with bulge formation from secular evolution (e.g. bar-induced inflows and bar buckling; see e.g. O’Neill & Dubinski 2003; Debatista et al. 2004; Mayer & Wadsley 2004; Athanassoula 2005, and references therein) appear to dominate the bulge population at mass ratios of similar amplitude ($B/T < 0.1 < 0.2$; see Kuijken & Merrifield 1995; Jogee et al. 2004; Kormendy & Kennicutt 2004; Fisher 2006; Fisher & Drory 2008; Weinzierl et al. 2009). Suggestively, this also corresponds to typical amplitudes of observed strong bars (references above and Eskridge et al. 2000; Laurikainen, Salo & Rautiainen 2002; Sheth et al. 2003; Marinova & Jogee 2007; Barazza et al. 2009).

Of course, real systems exhibit more complex behaviour then the simple scalings we derive here. Ultimately, detailed progress in modelling the interplay between continuous accretion of new substructure and cosmological driving of perturbations coupled to non-linear modes in galactic discs will require high-resolution N-body and hydrodynamic cosmological simulations. Some progress has begun towards modellling these processes in a proper cosmological context (see e.g. Bournaud & Combes 2002; Berentzen & Shlosman 2006; Gauthier et al. 2006; Governato et al. 2007; Kaufmann et al. 2007; Foyle et al. 2008; Kazantzidis et al. 2008; Romano-Diaz et al. 2008) – these studies highlight a key point here, that in a large regime of parameter space it is difficult to disentangle ‘secular’ and cosmological processes. Our goal here is not to derive a rigorous quantitative description of one or the other. However, the simple arguments here should help to constrain and focus the discussion of where and when (in realistic cosmological settings) ‘isolated’ evolution is important.

ACKNOWLEDGMENTS

We thank Lars Hernquist and T. J. Cox for helpful discussions, as well as Simon White and the anonymous referee for suggestions that greatly improved this manuscript. Support for PFH was provided by the Miller Institute for Basic Research in Science, University of California, Berkeley. DK acknowledges the support of the ITC fellowship at the Harvard College Observatory.

REFERENCES

Adams F. C., Ruden S. P., Shu F. H., 1989, ApJ, 347, 959
Akiyama M., Minowa Y., Kobayashi N., Ohta K., Ando M., Iwata I., 2008, ApJS, 175, 1
Athanassoula E., 2002a, ApJ, 569, L83

© 2009 The Authors. Journal compilation © 2009 RAS, MNRAS 401, 1131–1140
Stewart K. R., Bullock J. S., Barton E. J., Wechsler R. H., 2009b, ApJ, 702, 1005
Taylor J. E., Babul A., 2004, MNRAS, 348, 811
Toft S. et al., 2007, ApJ, 671, 285
Toomre A., 1981, in Fall S. M., Lynden-Bell D., eds, Structure and Evolution of Normal Galaxies. Cambridge Univ. Press, New York, p. 111
Trujillo I. et al., 2006, ApJ, 650, 18
van Starkenburg L., van der Werf P. P., Franx M., Labbê I., Rudnick G., Wuyts S., 2008, A&A, 488, 99
Velazquez H., White S. D. M., 1999, MNRAS, 304, 254
Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52
Weinberg M. D., 1985, MNRAS, 213, 451
Weinberg M. D., Katz N., 2007a, MNRAS, 375, 425
Weinberg M. D., Katz N., 2007b, MNRAS, 375, 460
Weinzierl T., Jogee S., Khochar S., Burkert A., Kormendy J., 2009, ApJ, 696, 411
Wetzel A. R., Cohn J. D., White M., 2009, MNRAS, 395, 1376
Woods D. F., Geller M. J., 2007, AJ, 134, 527
Woods D. F., Geller M. J., Barton E. J., 2006, AJ, 132, 197
Wyse R. F. G., Gilmore G., Norris J. E., Wilkinson M. I., Kleyba J. T., Koch A., Evans N. W., Grebel E. K., 2006, ApJ, 639, L13
Zemp M., Diemand J., Kuhlen M., Madau P., Moore B., Potter D., Stadel J., Widrow L., 2009, MNRAS, 394, 641

This paper has been typeset from a LaTeX file prepared by the author.