Analytic and numerical realizations of a disc galaxy

M. J. Stringer,[1,3]★ A. M. Brooks,[2,3] A. J. Benson[2,3] and F. Governato[4]

[1]*Department of Physics, South Road, Durham DH1 3LE*
[2]*Theoretical Astrophysics, Caltech, 1200 E. California Blvd., Pasadena, CA 91125, USA*
[3]*Kavli Institute of Theoretical Physics, Santa Barbara, CA 93106, USA*
[4]*Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA*

ABSTRACT

Recent focus on the importance of cold, unshocked gas accretion in galaxy formation – not explicitly included in semi-analytic studies – motivates the following detailed comparison between two inherently different modelling techniques: direct hydrodynamical simulation and semi-analytic modelling. By analysing the physical assumptions built into the GASOLINE simulation, formulae for the emergent behaviour are derived which allow immediate and accurate translation of these assumptions to the GALFORM semi-analytic model. The simulated halo merger history is then extracted and evolved using these equivalent equations, predicting a strikingly similar galactic system. This exercise demonstrates that it is the initial conditions and physical assumptions which are responsible for the predicted evolution, not the choice of modelling technique. On this level playing field, a *previously published* GALFORM model is applied (including additional physics such as chemical enrichment and feedback from active galactic nuclei) which leads to starkly different predictions.

Key words: methods: analytical – methods: numerical – galaxies: evolution – galaxies: formation.

1 INTRODUCTION

Two different approaches are widely used to test theories of galaxy formation. Both make use of developing computational resources to integrate a set of coupled differential equations forward in time, where each equation applies physical constraints, or empirical laws, to some of the system's properties.

The semi-analytic method (White & Frenk 1991) can be characterized by the intent to encapsulate as much of the physical behaviour as possible in the equations themselves, chosen to describe *systematic* properties (virialization, conservation of angular momentum, etc.) and thus leaving the minimum to numerical calculation.

Inherently different to this is the approach taken by simulations, where the ethos is to apply equations which are known to govern the *particular* properties (e.g. the gravitational force between two particles) and demonstrate, by skilfully programmed numerical calculation, that these reproduce macroscopic properties.

In practice, adjustment will always need to be made to account for the fact that each simulated particle represents many, many real particles. In traditional simulations which considered only gravity, this adjustment is purely mathematical. When gravity is the dominant interaction, results are therefore very robust and simulations are mostly accepted as being a faithful realization of the

Universe at large scales, where this is the case. A notable measure of this acceptance is that semi-analytic models have taken to using samples of haloes taken directly from simulations (Kauffmann et al. 1999; Hatton et al. 2003; Helly et al. 2003), rather than samples created statistically (Bower et al. 1993; Lacey & Cole 1993) using methods based on the analytic arguments of Press & Schechter (1974).

In general, though, there must come a scale below which the governing equations are no longer describing particle interactions but emergent phenomena. At these smaller scales, the two largely contrasting approaches begin to look more similar. As an immediate example, simulations which encompass an entire galaxy have not attempted to resolve individual stars and must therefore apply an equation which approximates the conversion rate of large bodies of gas into stars, based on locally averaged density. An equivalent and sometimes identical equation appears in semi-analytic models (Section 3.7).

Both approaches have had some success in *accounting for* the observed properties of galaxies. *Predictive* power, on the other hand, relies on surety in the underlying equations, or a clear understanding of how differing underlying assumptions map on to final observable quantities.

To bolster this crucial understanding, it is extremely helpful to study, in detail, the predicted evolution of a single, large galaxy and its satellites, as followed by these two different techniques. If the analytic model begins from the precise halo merger history found in the simulation, and applies the same assumptions about the physics

★E-mail: martin.stringer@durham.ac.uk

at smaller scales, will it predict the same observable properties and, if not, where and why do the two calculations diverge?

1.1 Previous research

Existing quantitative comparisons between simulations and semi-analytic models have focused on *samples* of galaxies. The approach which is adopted here – adapting the semi-analytic model to emulate the assumptions made by the simulation – was also used by Helly et al. (2003) to demonstrate that the cooled mass predicted by the two approaches can be consistent. In particular, consistency was demonstrated over two orders of magnitude in halo mass. Earlier works by Benson et al. (2001) and Yoshida et al. (2002), using similar techniques, reached broadly similar conclusions.

A broader comparison was made by Cattaneo et al. (2007) of two predicted galaxy populations: one by an smoothed particle hydrodynamics (SPH) model and the other by version of the GalICS semi-analytic model, modified to emulate the SPH method. Again, consistency is demonstrated between the two techniques once the semi-analytic model is appropriately 'stripped down', but this consistency comes at the expense of agreement with observational constraints, which has been met by the full version of GalICS.

1.2 Outline

To extend this previous research, in this present work, we will compare in detail the formation of a single galaxy in a GASOLINE numerical simulation and in the GALFORM semi-analytic model. The structure of this paper is as follows. Section 2 is a review of the broader aspects of the GASOLINE code and of this particular simulation. Section 3 offers a detailed analysis of the way that GASOLINE models, each of the following key physical processes, together with an explanation of how the particular approach in question can be emulated in the GALFORM model: Section 3.1, Merging; Section 3.2, Accretion; Section 3.3, UV heating; Section 3.4, Gas distribution; Section 3.5, Cooling; Section 3.6, Disc formation; Section 3.7, Star formation; Section 3.8, Feedback; Section 3.9, Disc stability. Particularly, relevant aspects of the GALFORM model are reviewed (or referenced) when appropriate. General aspects of the model can be found in Cole et al. (2000) and Bower et al. (2006). In Section 4.1, by applying these same underlying physical assumptions to GALFORM, the two modelling techniques are shown to make broadly consistent predictions. In Section 4.2, having demonstrated this consistency, GALFORM is then used to find how this halo would evolve under the physical assumptions made in previous studies (which have been shown to successfully match the *collective* properties of galaxies). Section 5 provides the summary.

2 GASOLINE

The 'MW1' simulation analysed in this paper was generated using GASOLINE, an N-Body + SPH Parallel Treecode (Wadsley, Stadel & Quinn 2004), within a flat, Λ-dominated *Wilkinson Microwave Anisotropy Probe* 1 (*WMAP*1) cosmology ($\Omega_0 = 0.3, \Lambda = 0.7, h = 0.7, \sigma_8 = 0.9, \Omega_b = 0.039$). A moderate resolution version of this run was originally discussed in detail in Governato et al. (2007), and the higher resolution version analysed here has subsequently been used for a number of studies (Read et al. 2009; Zolotov et al. 2009). High-mass and force resolution (see Table 1) was achieved using the volume renormalization technique (Katz & White 1993).

Table 1. Resolution details for the GASOLINE simulation.

DM particles	$7.6 \times 10^5 \, M_\odot$
Gas particles	$1.3 \times 10^5 \, M_\odot$
Star particles	$3.8 \times 10^4 \, M_\odot$
Force resolution	0.3 kpc

This simulated galaxy matches the observed mass–metallicity relationship at varying redshifts (Brooks et al. 2007; Maiolino et al. 2008). This success demonstrates that these simulations overcome the historic 'overcooling problem' (Mayer, Governato & Kaufmann 2008), in which gas in simulations cools rapidly, forming stars too quickly and early, and thus producing a mass–metallicity relation that is too enriched at a given stellar mass. Brooks et al. (2007) showed that reproducing the mass–metallicity relationship was a result of the adopted feedback prescription. This feedback mechanism, combined with high numerical resolution,[1] avoids spurious loss of angular momentum. Furthermore, the simulated galaxies also match the observed Tully–Fisher relationship (Governato, Mayer & Brook 2008; Mayer et al. 2008; Governato et al. 2009).

GASOLINE's star formation and supernovae feedback scheme is described in detail in Stinson et al. (2006), and parameters from Governato et al. (2007) are adopted for MW1. The star formation recipe reproduces the Kennicutt–Schmidt law, and each star particle is born with a Kroupa initial mass function (Kroupa, Tout & Gilmore 1993). Combined with a prescription for a cosmic UV background (Haardt & Madau 1996) that mimics reionization and unbinds non-collapsed baryons from simulated haloes with total masses below about $10^9 \, M_\odot$, this feedback scheme drastically reduces the number of luminous satellites containing a significant mass of stars, avoiding the well-known 'substructure problem' (White & Frenk 1991; Kauffmann, White & Guiderdoni 1993; Quinn, Katz & Efstathiou 1996; Moore et al. 1999).

As described in Stinson et al. (2006), supernova energy is deposited into the ISM mimicking the blast-wave phase of a supernova (following the Sedov–Taylor solution). Cooling is turned off for gas particles within the supernova blastwave radius for the duration that the remnant is expected to adiabatically expand (Section 3.8). Supernova feedback regulates star formation efficiency as a function of halo mass, resulting in the mass–metallicity relationship described above. This regulation of star formation also leads to realistic trends in gas fractions, with the lowest mass galaxies being the most gas rich (Brooks et al. 2007), and reproducing the observed incidence rate and column densities of Dampled Lyman α systems at $z = 3$ (Pontzen et al. 2008).

MW1 and its most massive progenitor halo were identified at each simulation output step using AHF[2] (Gill, Knebe & Gibson 2004; Knollmann & Knebe 2009). AHF determines the virial radius, r_v, for each halo at each output time-step using the overdensity criterion for a flat universe following Gross (1997). The full gas accretion history of MW1 is described in detail in Brooks et al. (2009), and used in this study for comparison to GALFORM.

---

[1] With $\sim 4.8 \times 10^6$ total particles within the virial radius at $z = 0$.
[2] AMIGA's Halo Finder, available for download at http://popia.ft.uam.es/AMIGA.

3 TRANSLATION FROM SIMULATION TO SEMI-ANALYTICS

3.1 The merger tree

The common point between the two different models of galaxy formation is the halo merger tree. This is generated by finding bound structures in the 'MW1' simulation of Brooks et al. (2009) using AHF. A virial mass is assigned to each structure and its fate at each time-step is categorized, being either

    (i) isolated,
    (ii) inside the virial radius of a larger system or
    (iii) merged.

Two criteria are used to classify a galaxy as 'merged':

    (i) if the dark matter (DM) mass decreases by more than 50 per cent in a single time-step or
    (ii) if the number of DM particles falls below 64.

These are both significantly different from the definition adopted by GALFORM (see Cole et al. 2000), so a detailed comparison of merger times, along the lines of Jiang et al. (2008), is deferred for another study.

When a structure changes from category 1 to category 2, the GALFORM code updates its status to a 'satellite' and predicts its subsequent evolution based on its approach velocity v and position r at this juncture. These two properties are combined into the dimensionless parameter Θ (Section A1), which characterizes the orbit. The distribution of values for Θ in this simulation was found to be consistent with the most recent assumption applied in GALFORM (Fig. 1).

This orbital information is then used to estimate a merger time based on the standard Chandrasekhar formula (A6) as defined by Lacey & Cole (1993). Details of this part of the model can be found in Appendix A and in Cole et al. (2000; section 4.3).

The resulting dark and baryonic components are shown in Fig. 2. Being generated purely from the derived merger tree, the two realizations might be expected to be identical. One reason this is not quite the case is because the total virial mass, M_v, is measured directly, but the relative composition of baryons and DM may vary a little from halo to halo, and from one realization to the other.
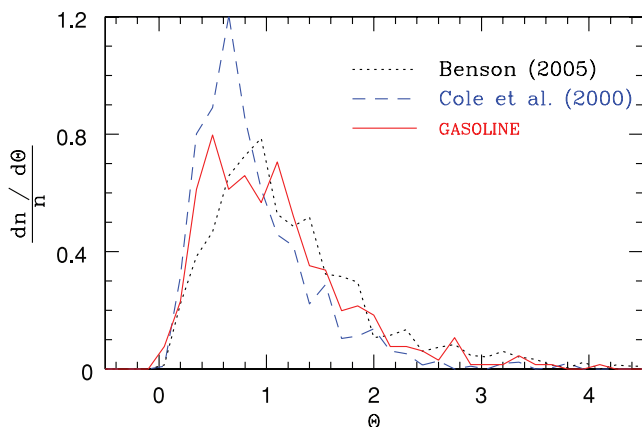

Figure 1. Orbital parameters randomly selected from the distribution (A5) of Cole et al. (2000) are shown in the dashed line, while the dotted line shows the distribution from Benson (2005). The solid line shows the distribution found in the simulation of Brooks et al. (2009) [which results from applying equation (A4) to the positions and velocities of the satellites].
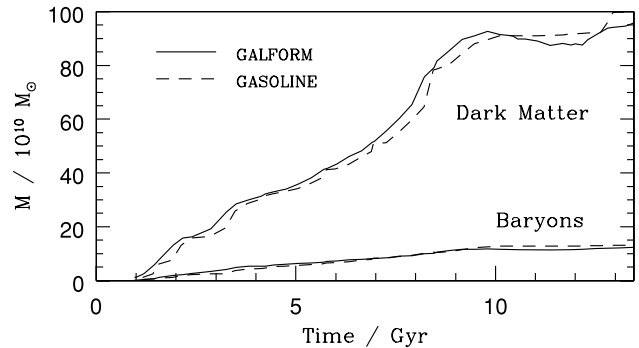

Figure 2. Confirmation that the mass evolution of the 'MW1' system which emerged in the GASOLINE simulation of Brooks et al. (2009) (solid lines) matches that contained in the extracted merger history used by GALFORM (dashed lines).

3.2 Accretion

The temperature history of accreted gas was investigated by Kereš et al. (2005) in a simulation containing 1120 galaxies. The distribution of gas particles in terms of their maximum past temperature is bimodal, suggesting that classification into hot and cold modes of accretion is appropriate. The contribution from each of these modes is then found to depend strongly on halo mass and on environment. Notably, the halo mass ($M_v \sim 3 \times 10^{11}$ $M_\odot$), which separates low-mass galaxies dominated by cold accretion and higher mass galaxies dominated by the hot mode, is close to that found by analytic arguments (Birnboim & Dekel 2003).

The decomposition of the accreting gas into different modes is a principle feature of the analysis by Brooks et al. (2009). The three modes of accretion are defined as follows (along with their analogues in the semi-analytic realization).

(i) *Clumpy accretion* is defined to include 'any particles that have, at any previous output step, belonged to a galaxy halo other than the main halo we are considering' (Brooks et al. 2009).

The information needed to evaluate this component in the semi-analytic model is readily available from the merger tree, being the virial masses, M_v, of the existing, progenitor haloes:

$$M_{\text{clumpy}}(\text{new}) = \frac{\Omega_b}{\Omega_M} \sum_{\text{progenitors}} M_v(\text{progenitor}). \tag{1}$$

The subsequent fate of any baryons within these substructures is decided by their inbound trajectory and considerations of dynamical friction, as described in Section 3.1 and Appendix A.

(ii) *Smooth accretion: shocked component.* The following criteria are required for a gas particle in the simulation to be categorized as shocked:

$$(a) \ T \geq \frac{3}{8}\mathcal{T}_v \quad \text{and} \quad (b) \ \Delta\rho > 4\left(\frac{3\mathcal{T}_v}{3T}\right)^{\frac{3}{2}} - 1\,, \tag{2}$$

where $\mathcal{T}_v$ is the virial temperature defined in GASOLINE,[3]

$$k_B\mathcal{T}_v \equiv \frac{GM_v\mu m_H}{3r_v}, \tag{3}$$

and $\Delta\rho$ is the fractional increase in density since the previous time-step. The latter is an entropy criterion and the factor of 4 is motivated by the Rankine–Huginot shock jump conditions for a singular isothermal sphere.

---

[3] The virial temperature in GALFORM is defined differently, being a factor of 3/2 higher, $T_v \equiv \frac{3}{2}\mathcal{T}_v$, so (2a) corresponds to $T \geq \frac{9}{16}T_v$.

Though low resolution can artificially broaden shocks, Brooks et al. (2009) verify that this run is of sufficiently high resolution to resolve shock discontinuities. On the more general issue, of the agreement between cooling rates in reality and values calculated using SPH, the reader is referred to Wadsley et al. (2004).

Within the framework of GALFORM, finding an appropriate analogue to this condition is difficult. However, one relevant calculation that *is* made compares the cooling time-scale for gas at temperature T:

$$t_{cool}(r) = \frac{3\mu m_H k_B T}{2\rho_{gas}(r)\Lambda(T, Z)} , \qquad (4)$$

and the free-fall time from a radius, r, enclosing mean density $\bar{\rho}$

$$t_{ff}(r) = \sqrt{\frac{3}{8\pi G \bar{\rho}(r)}} = \frac{r_v}{2}\sqrt{\frac{\mu m_H}{k_B T_v}}. \qquad (5)$$

The cooling function Λ is found for given temperature T and metallicity Z, by reference to the published tables of Sutherland & Dopita (1993). ρ_{gas} is the gas density, $\bar{\rho}(r)$ the mean density of all the mass enclosed by radius r and μm_H is the mean particle mass.

So, in the context of GALFORM accreting gas is designated as shocked if

$$t_{cool}(T_v, r_v) > t_{ff}(T_v, r_v) , \qquad (6)$$

this being a reasonable, simple equivalent to (2a).

(iii) *Smooth accretion: unshocked component.* Gas in the simulation which fails to meet either of the criteria in (2) is categorized as unshocked. Similarly, accreting gas in GALFORM is designated as unshocked if $t_{cool} < t_{ff}$.

The relative contributions of these three components are shown in Fig. 3 for both realizations. Being determined directly from the merger tree, it is unsurprising that the clumpy components are quite consistent.

That the shocked component is overestimated by the semi-analytic model is entirely as anticipated by Brooks et al. (2009), though the eventual impact this has on the predicted properties of the galaxy may not be quite as severe, as will be seen in the forthcoming sections.

3.2.1 Alternative criterion

Birnboim & Dekel (2003) investigate the conditions for the existence of virial shocks in galactic haloes, culminating in the following expression for the existence of a shock:

$$\frac{\rho_{gas}(r_v)r_v\Lambda(T_v)}{4(\mu m_H)^2 |u_0|^3} < 0.0126, \qquad \left[u_0^2 = \frac{16k_B T_v}{3\mu m_H}\right]. \qquad (7)$$

Referring to equations (3)–(5) above, this condition can be rewritten as

$$t_{ff} < 0.2 t_{cool}. \qquad (8)$$

As can be seen from Fig. 3, the cooling time does not exceed the free-fall time by a factor of 5 until about 8 Gyr. The criterion (8) would therefore imply that all smoothly accreted gas will be unshocked before that point, corresponding to an unshocked:shocked ratio of 2 or 3, as opposed to 1/2 in gasoline or 1/6 according to (6).

So, for this particular mass aggregation history, one simple estimate (6), overestimates the shocked component in the simulation and another, more thorough estimate (7), gives an underestimate. This suggests that the error in assuming spherical symmetry may unfortunately exceed any accuracy that can be gained by analysis of shock conditions. Calibrating some of these analytic models
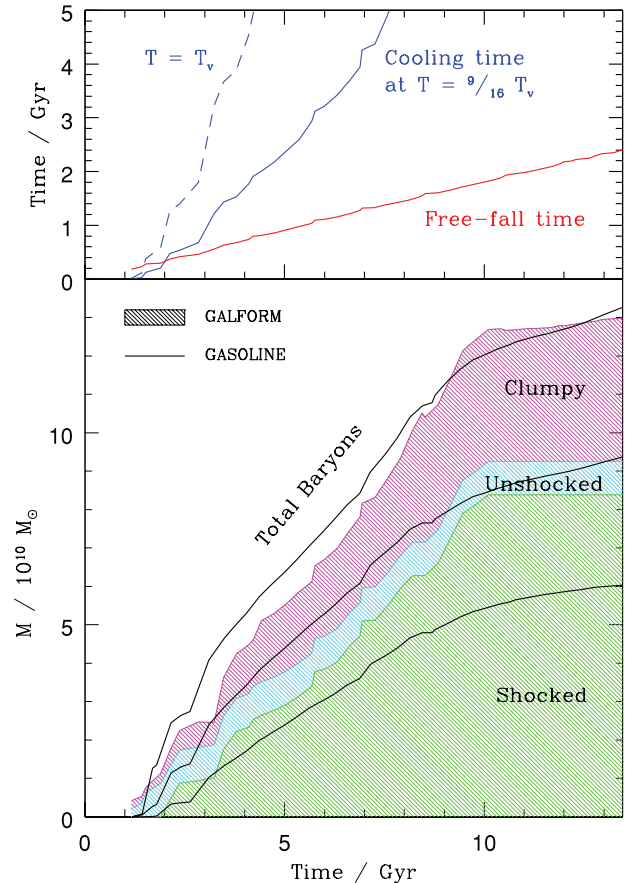

Figure 3. An analysis of accretion on to the MW1 galaxy in GASOLINE, in the light of assumptions associated with GALFORM. The top panel shows the cooling time for a gas particle at the virial radius (4) at two relevant temperatures, together with the time-scale for free-fall under gravity (5). In the lower panel, the solid lines show how accretion is divided amongst these three modes in the simulation. The shaded areas show how the equivalent components would be estimated in GALFORM. As a simple attempt to emulate (2a), shocked and unshocked are distinguished by comparing the cooling time at $T = {}^9/_{16}T_v$ to the free-fall time (as explained in Section 3.2b).

to simulations may help quantitatively account for the complex geometry.

3.3 UV heating

A UV background is incorporated by Brooks et al. (2009) which is based on an updated model by Haardt & Madau (1996). This background 'turns on at $z = 6$' and heats the gas in haloes so that only those with '100 or more DM particles are of sufficient mass to retain bound gas particles'. For the resolution in this case, 100 DM particles corresponds to $M_{cut} = 7.6 \times 10^7 \, M_\odot$.

The GALFORM model is built with two corresponding parameters: a redshift, z_{cut}, below which the UV background heating is effective (set to $z_{cut} = 6$ for consistency with the above) and a halo velocity, v_{cut}, below which all cooling is suppressed. For a virial overdensity, $\Delta_v = 300$, the mass cut-off, M_{cut}, mentioned above translates to a cut-off halo velocity of

$$v_{cut} = (\Delta_v/2)^{\frac{1}{6}} (1 + z)^{\frac{1}{2}} (\Omega_M G M_{cut} H_0)^{\frac{1}{3}} \approx 30 \, \text{km s}^{-1}. \qquad (9)$$

These values are duly applied, but one key difference between the modelling of this process remains. The UV background in the simulation can *slow* cooling in larger haloes without necessarily

completely suppressing it, while in GALFORM suppression is all or nothing (below and above v_{cut}, respectively).

3.4 Hot gas distribution

The distribution of hot gas that is assumed in the GALFORM model was recently updated, in the light of the work by Sharma & Steinmetz (2005), to the following density profile for hot gas of mass M_H:

$$\rho(r) = \frac{\rho(0)}{(1 + r/r_H)^3}, \quad \left[\rho(0) \equiv \frac{M_H/4\pi r_H^3}{\int_0^{r_v/r_H} x^2(1+x)^{-3}\,\mathrm{d}x}\right]. \quad (10)$$

This form was assumed in recent work on disc formation and structure (Stringer & Benson 2007), and will also be used here.

The reason for this choice is clear from Fig. 4, which demonstrates that it is a satisfactory description of the distribution of diffuse gas in the outer parts of this simulated system. In this figure, the analytic form for the hot halo gas (10) and cold disc gas (20) have been fitted to the simulation results, yielding the respective scalelengths:

(i) halo core radius, r_H,
(ii) Disc scalelength, r_D.

(with the initial constraint on each line that it integrates to the correct total mass for that component.) These scalelengths, found from fitting the simulation results, can be scrutinized against those deduced from the usual analytic assumptions.

First, the halo core radius is typically assumed to be a constant fraction of the virial radius, which is not quite the case in this particular simulation (Fig. 4). However, in the absence of any obvious pattern in the evolution of r_H, the usual simple scaling is applied, with the choice

$$r_H = r_v/3 \quad (11)$$

being the best approximation. Fig. 5 compares this to values found from fitting to the true gas distributions in Fig. 4, the result looking favourable. This straightforward scaling with the virial radius (11) will therefore be used in the matched GALFORM realization. The cold gas scalelengths, also shown in Fig. 5, are not inputs to GALFORM, but are calculated within the model as described in Section 3.6.

To understand the relevance of the hot gas distribution on the system's evolution as a whole, it is worth reviewing the precise treatment of gas cooling in GALFORM.
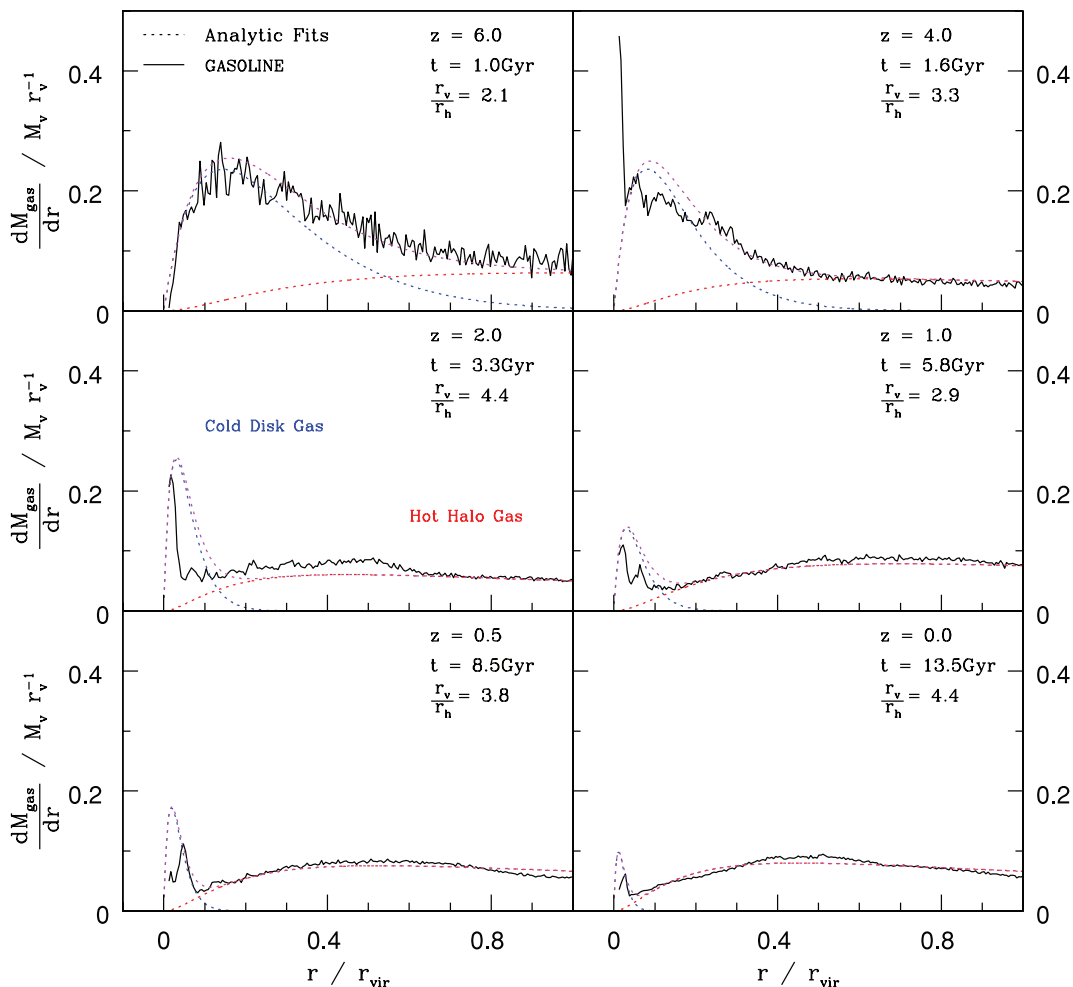

Figure 4. The total gas density profile from the simulation of Brooks et al. (2009) at six different redshifts, shown as solid lines. The dotted lines show the analytic forms which are assumed by the GALFORM model for the disc gas (equation 20, blue line) and diffuse halo gas (equation 10, red line). The sum total of these two components is also shown (magenta line). Each of these depends on a single parameter, fitted respectively by the disc scalelength, r_D and the core radius of the hot gas profile, r_H. (The integral under each analytic component is already fixed to match the total hot and cold gas mass within the virial radius.) Because there are so many more points to fit the hot component, these are much better matches than the cold gas. As it is the hot scalelength (*not* the cold gas scalelength) which is actually an input to GALFORM, this convenient bias has not been corrected.
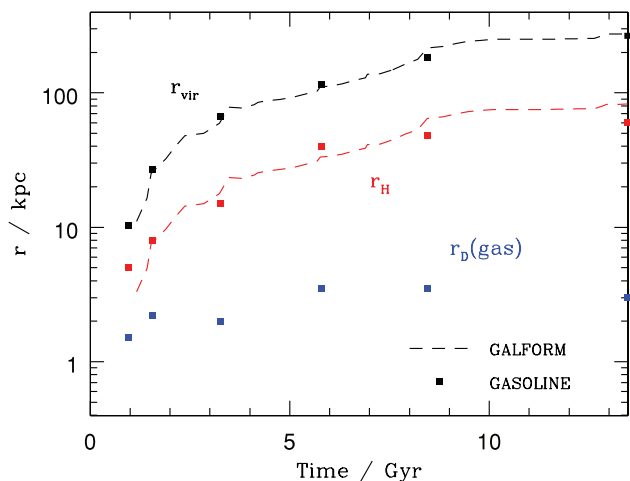
Figure 5. The evolution of three key scalelengths of the system: virial radius, r_v, halo gas scalelength, r_H (10), and disc scalelength, r_D (20). Dashed lines show the values that will be assumed by GALFORM. Points show the values which best fit the gas profile of the simulation (Fig. 4).

3.5 Cooling

The cooling model described by Cole et al. (2000) determines the mass of gas able to cool in any time-step by following the propagation of the cooling radius in a notional hot gas density profile.[4] This profile is fixed when a halo is flagged as 'forming' and is only updated when the halo undergoes another formation event. The mass of gas able to cool in any given time-step is equal to the mass of gas in this notional profile between the cooling radius at the present step and that at the previous step. The cooling time is assumed to be the time since the formation event of the halo. Any gas which is reheated into or accreted by the halo is ignored until the next formation event, at which point it is added to the hot gas profile of the newly formed halo. The notional profile is constructed using the properties (e.g. scale radius, virial temperature, etc.) of the halo at the formation event and retains a fixed metallicity throughout, corresponding to the metallicity of the hot gas in the halo at the formation event.

This work makes use of a new cooling model, which will be described in full detail in Benson & Bower (2010). Rather than arbitrary 'formation' events, this model uses a continuously updating estimate of cooling time and halo properties. The properties described in Section 3.4 (density normalization, core radius) are reset at each time-step. The previous infall radius[5] (i.e. the radius within which gas was allowed to infall and accrete on to the galaxy) is computed by finding the radius which encloses the mass previously removed from the hot component in the current notional profile. The new model must supply an alternative estimate of the time available for cooling in the halo, t_{avail}, from which the cooling radius can be computed in the usual way (i.e. by finding the radius in the notional profile at which $t_{cool} = t_{avail}$). This is done by considering the energy radiation rate *per particle*, $\dot{\epsilon}_r$ and the thermal energy *per particle*,

ϵ_{th}, which are estimated by making standard assumptions:

$$\epsilon_{th} = \frac{3}{2}k_B T_v \qquad \text{and} \qquad \dot{\epsilon}_r = \Lambda(T_v, Z)n. \qquad (12)$$

In terms of these quantities, the cooling time (4) is simply

$$t_{cool}(r, t) = \frac{\epsilon_{th}(t)}{\dot{\epsilon}_r(r_{cool}, t)}. \qquad (13)$$

At any time, the model needs to identify some radius in the hot halo, r_{cool}, where the gas has had just enough time to radiate all its thermal energy. This *as yet undetermined* radius is defined to satisfy the condition that the cooling time for gas particles at this point is equal to the time available for them to cool, which is estimated in terms of the mean energy radiation rate per particle,[6] $\overline{\dot{\epsilon}}_r$:

$$t_{avail}(t) = \frac{\int_0^t \overline{\dot{\epsilon}}_r(t_i)dt_i}{\overline{\dot{\epsilon}}_r(t)}. \qquad (15)$$

Equating (13) and (15) then gives

$$t_{cool} = t_{avail},$$

$$\frac{\frac{3}{2}k_B T_v}{\Lambda(t)n(r_{cool}, t)} = \frac{\int_0^t \Lambda(t_i)\overline{n}(t_i)dt_i}{\Lambda(t)\overline{n}(t)}, \qquad (16)$$

where $\bar{n}$ is the mean number density inside the virial radius.

This approach must account for the hierarchical assembly of a halo, so the denominator in the right-hand side strictly includes a sum over all progenitors, i, of this system which exist at time t_i. Incorporating this sum into (16), and rearranging, then yields a final expression to be solved for the cut-off density, $\rho_{cool} \equiv \rho(r_{cool})$:

$$\frac{\rho_{cool}(t)}{\overline{\rho}(t)} = \frac{\frac{3}{2}k_B T_v}{\int_0^t \left[\sum_i \Lambda_i(t_i)\overline{n}_i(t_i)\right] dt_i}. \qquad (17)$$

This equation is notably independent of halo structure, a factor which only comes in when the limiting density is used to compute the total cooled mass. The scalelength, r_H, appears in the normalization of the hot gas profile (10), and hence effects the mass *enclosed* by the cut-off density, ρ_{cool}.

This can be appreciated graphically. Solutions to (17) are illustrated in Fig. 6 for three choices of r_H and three examples of cut-off density. For the higher cut-off density, a smaller r_H (more centrally concentrated) will lead to a greater cooled mass. For the case of a low cut-off density, a centrally concentrated profile will actually lead to a lesser cooled mass, contrary, perhaps, to intuition.

3.6 Disc formation

In GALFORM, galactic discs are assumed to form from the cooled gas, conserving the net angular momentum that it possessed when it was part of the halo. The assumed distribution for the specific angular momentum, j, of the hot halo gas is taken from Sharma & Steinmetz (2005):

$$\frac{dM(j)}{dj} \propto j^{\alpha-1} \exp\left(-\frac{\alpha j}{\langle j \rangle}\right). \qquad (18)$$

---

[4] We refer to this as a 'notional' profile since it is taken to represent the profile before any cooling can occur. Once some cooling occurs, presumably the actual profile adjusts in some way to respond to this and so will no longer look like the notional profile, even outside of the cooling radius.
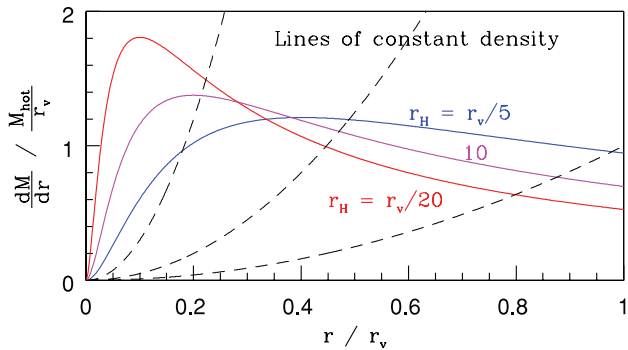[5] Defined as the lesser of the cooling and free-fall radii.

[6] This is proportional to the mean number density $\bar{n}$:

$$\overline{\dot{\epsilon}}_r \equiv \frac{\Lambda \int_0^N n(N')dN'}{N} = \alpha(r_H)\Lambda\bar{n} \qquad (14)$$

The factor α does depend on halo structure (i.e. on r_H) but, if the distribution of hot gas *relative to the virial radius* is the same for each halo, as implied in this case by (11), then this factor cancels out in (16).

Figure 6. Three normalized hot gas distributions (solid lines), corresponding to three different values of the scalelength r_H applied to the profile (10). Intersections between a solid line and one of the dotted lines correspond to solutions to equation (17) for a particular value of r_H and a particular cut-off density. The cooled fraction, M_{cool}/M_{hot}, is given by the integral under the solid curve up to the intersection with the relevant cut-off density.

The value of α was found to lie between 0.5 and 1.5, but the mean value $\alpha = 0.89$ is adopted here. This is normalized to enclose the correct total mass and the mean specific angular momentum of the hot gas, $\langle j \rangle$, which is found directly from the spin parameter:

$$\lambda \equiv \frac{\langle j \rangle |E|^{1/2}}{GM_v^{3/2}}, \qquad \left[\langle j \rangle = \frac{1}{M_v} \int_0^{M_v} j \, \mathrm{d}M(j) \right]. \quad (19)$$

Details of the calculation of the halo energy, E, can be found in Cole et al. (2000; section 3.2.1). The disc scalelength, r_D, is recalculated after each time-step such that angular momentum is conserved in the system. The calculation is based on the following surface density distribution (applied to each disc component, i):

$$\Sigma_i(R) = \Sigma_i(0) \, \mathrm{e}^{-\frac{r}{r_D}}, \qquad \left[\text{so } M_i = 2\pi \Sigma_i(0) r_D^2 \right] \quad (20)$$

and a circular speed, v_c, which is constant with radius. The subsequent evolution of the disc is then followed by applying physical assumptions which are as close as possible to those made in the simulation, as is explained in the following sections.

As with the subhalo trajectories, the parameter λ can ordinarily be assigned at random, from a specified distribution. In this case, its value is measured for each of the simulated haloes and passed on to GALFORM, along with the rest of the merger tree information (Section 3.1).

The goal here is not to prove that parameters like λ, and Θ must necessarily be set by simulations in order to validate the results of any semi-analytic model. Rather, it is an example of the correspondence between these initial conditions and the properties of the galaxy which they produce. This should reassure theorists that, when such models are applied to large samples of galaxies, their individual properties will be representative as long as these governing, random, but *physical* parameters are drawn from the correct natural distributions.

3.7 Star formation

At face value, GASOLINE and GALFORM would appear to be applying the same assumptions regarding the conversion of cold gas to stars. These respective assumptions can be found from equation (4) of Stinson et al. (2006) and equations (4.4) and (4.14) of Cole et al.
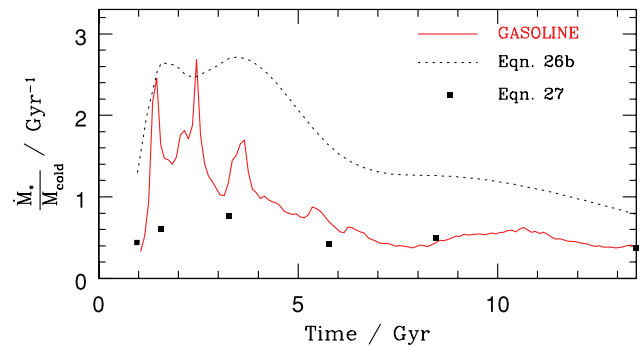

Figure 7. The star formation efficiency from the simulation is shown as a solid line. The points show the result of applying the corrected star formation rate (24) to the disc scalelengths, r_D which best fit the simulation (Fig. 4). The illustrative approximation of a constant scaleheight, $h = 200$ pc, is used. The faint dotted line shows the efficiency found in empirical studies, which is given by equation (21b) with $\epsilon_\star = 0.017$ (Kennicutt 1998).

(2000):[7]

$$(a) \quad \frac{\mathrm{d}\rho_\star}{\mathrm{d}t} = c_\star \frac{\rho_{cold}}{t_{dyn}} \qquad \text{and} \qquad (b) \quad \psi = \epsilon_\star \frac{M_{cold}}{\tau_{disc}}, \quad (21)$$

where ψ is the total instantaneous star formation rate in the whole disc and $\rho_\star$ is the local value per unit volume. The time-scales which appear in these expressions have similar physical motivation, but are not identical. One, t_{dyn}, is the time for gravitational collapse of a spherically symmetric region of local density ρ_{cold}. The other, τ_{disc}, is the angular period of the whole system at the half-mass radius, $r_{\frac{1}{2}}$:

$$(a) \quad t_{dyn} \equiv \frac{1}{\sqrt{G\rho_{cold}}} \qquad \text{and} \qquad (b) \quad \tau_{disc} \equiv \frac{r_{1/2}}{v_c}. \quad (22)$$

So, despite the apparent similarity of the two expressions in (21), they can lead to quite different instantaneous *total* star formation rates, ψ. To illustrate this, one can consider the surface density distribution (20) which GALFORM assumes, together with some decrease in density away from the plane of the disc:

$$\rho_i(R, z) = \rho_i(R, 0) \, \mathrm{e}^{-\frac{|z|}{h}}, \qquad \left[\text{so } \Sigma_i(R) = 2\rho_i(R, 0)h \right]. \quad (23)$$

Applying the local star formation assumption of (21a) to this distribution gives

$$\psi_a = 2\pi \sqrt{G} \rho_{cold}(0, 0)^{\frac{3}{2}} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{3|z|}{2h}} \mathrm{d}z \int_0^{\infty} \mathrm{e}^{-\frac{3r}{2r_D}} R \, \mathrm{d}R$$

$$= c_\star \left(\frac{2}{3} \right)^3 M_{cold}^{3/2} \sqrt{\frac{G}{4\pi h r_D^2}}. \quad (24)$$

Comparison between this and the usual prescription in GALFORM exposes the big *physical* difference between the two star formation models; (21a) considers the gravitational collapse of the gas due to its own gravity while (21b) includes the gravity of all components,[8] including the stars and DM (M_{tot} rather than just M_{cold}).

The derived analytic equivalent (24) is shown in Fig. 7 (using the fitted values of r_D from Section 3.4 and the illustrative approximation[9] of a constant scaleheight, $h = 200$ pc) alongside the full

---

[7] An additional factor involving the circular speed appears in (4.11) of Cole et al. (2000). This is omitted here (equivalent to setting $\alpha_\star = 0$).

[8] The DM does influence the creation of stars due to its effect on the overall stability, but this is a separate part of the analytic treatment (Section 3.9).

[9] This choice is not too critical since h appears only to the one half power in equation (24).

simulation results. Discrepancies are larger at early times when the assumed cold gas density distribution is tenuous, but the match improves as the disc settles and is excellent at late times, showing that (24) is as consistent with the assumptions in GASOLINE as is possible within this particular analytic framework.

3.7.1 Burst star formation

During merger events, and if the disc is deemed to be unstable (Section 3.9), cold disc gas in the GALFORM model is assumed to funnel to the centre of the new system and be converted to stars within a time-step, which in this case is about 0.2 Gyr (Cole et al. 2000; Section 4.3). The incorporation of this non-equilibrium effect should broadly correct the total star formation rate on occasions when the assumed density profile (20) is not such a good description of the true gas distribution. This is relevant here in the early stages of the simulation (as apparent when comparing the points and solid line in Fig. 7).

3.7.2 Schmidt–Kennicutt Law

Equation (21b) can be recognized as a version of the established empirical relation between star formation rate and gas supply (Kennicutt 1998), for which the constant of proportionality has the value $\epsilon_\star \approx 0.017$. The corresponding star formation efficiency for this simulated system can be calculated from its approximate half-mass radius and circular velocity (Fig. 10) and this is shown in Fig. 7 as a guide.

3.8 Feedback

3.8.1 Heating of disc gas

The MW1 simulation treats each type II supernova event as a blast wave which expands to a maximum radius, r_{max}, and the enclosed volume is assumed to be unable to cool for as long as the shell can be expected to survive, t_{max}. If the star formation rate varies little over this time-scale, and the shell subsequently cools rapidly, this will correspond to a heated mass,

$$M_{heat}(t) \approx \frac{\dot{M}_\star t_{max}}{m_{SN}} \left(\frac{4\pi \rho_{cold} r_{max}^3}{3} \right) \equiv \dot{M}_\star(t) t_{heat} , \qquad (25)$$

where m_{SN} is the average mass of stars formed per supernova. Now, (25) conceals the fact that both r_{max} and t_{max} (and hence t_{heat}) are, rightly, functions of the local gas density, ρ and pressure, P. The precise dependence assumed is found in equations (9) and (11) of Stinson et al. (2006). Combining this into (25) gives

$$t_{heat} \sim \rho(\rho^{-0.16}P^{-0.2})^3 \rho^{0.34}P^{-0.7} \sim \rho^{0.86}P^{-1.3} . \qquad (26)$$

Referring to the phase diagram in fig. 4 of Stinson et al. (2006), gas in star-forming regions appears to be approximately isothermal at $T \approx 10^4$ K (which is expected due to the steep gradient in the cooling curve at this temperature). This leads to the suggestion that the environmental dependence in (26) will be minimal, the pressure and density terms almost cancelling for constant temperature. Though tenuous, this provides at least a rough explanation for the extremely simple dependence which does indeed emerge from the full calculation, as shown in Fig. 8:

$$M_{heat}(t) \approx \dot{M}_\star(t) t_{heat}, \qquad \left[t_{heat} \approx 7 \times 10^6 \text{ yr} \right] . \qquad (27)$$
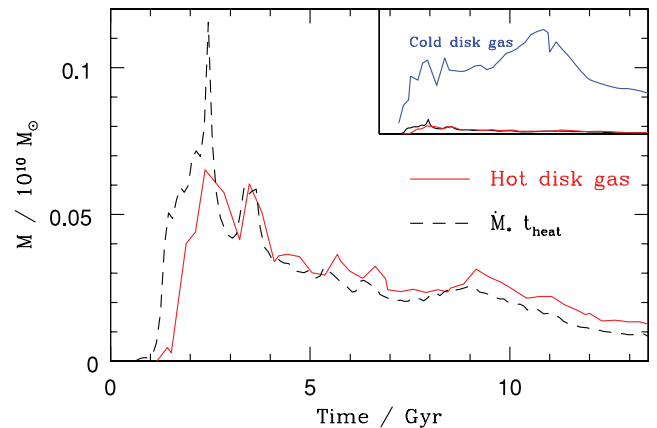

Figure 8. A simple illustration of the correlation between star formation rate and the mass of heated gas present in the disc. The effective time-scale from equation (27), $t_{heat} = 7 \times 10^6$ yr. The inset panel gives an indication of the almost negligible fraction of the total disc gas which this represents.

The precise definition of M_{heat} used in the figure is the mass of disc gas which fails to satisfy the temperature criterion for star formation ($T < 30\,000$ K). Thus, the integrated effect of feedback will be to slightly reduce the fraction of the total disc gas which is available for star formation:[10]

$$\dot{M}_\star = \frac{M_{cold} - M_{heat}}{\tau_\star} = \frac{M_{cold}}{\tau_\star + t_{heat}} , \qquad (28)$$

where $\tau_\star$ is the particular time-scale derived from equation (24). This revision is duly applied in GALFORM, being a fair reflection of the physical assumptions in the simulation.

3.8.2 Outflow

To discover the extent of outflow from the galaxy in the simulation, all particles were identified which have moved from the inside to the outside of a radius of 30 comoving kpc. The total mass of such material turned out to be even less than the heated disc mass (see Fig. 8) and can safely be ignored for the purposes of this study. The outflow of gas as a result of supernovae in the central galaxy is therefore set to zero in this version of GALFORM, for consistency with the GASOLINE simulation.

It is important to note that this feedback model is more effective at lower halo masses, due to shallower potential wells (Brooks et al. 2007). This dependence is also accounted for in the usual GALFORM feedback model, where the rate of mass outflow is assumed to be at least equal to the star formation rate, and thousands of times higher for smaller systems (Bower et al. 2006). The consequence of this difference is investigated in Section 4.3.

3.9 Disc stability

The analytic condition which is used to assess the stability of discs in the GALFORM model follows Efstathiou, Lake & Negroponte (1982).

---

[10] For a massive galaxy such as this, the star-forming time-scale is typically $\tau_\star > 10^8$ yr, so the correction to allow for reheating actually makes very little difference. Quick examination of the inset to Fig. 8 confirms this; the fraction of the total gas which gets heated above the star-forming limit is tiny.
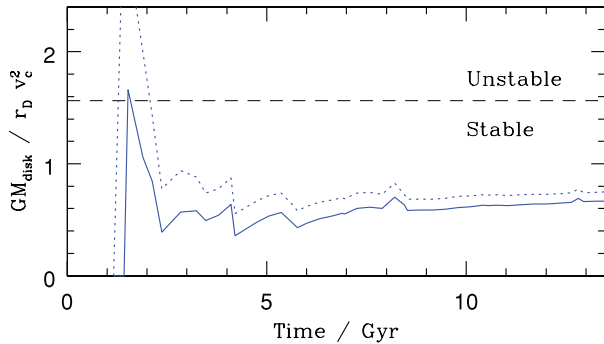
Figure 9. The dimensionless ratio which is used as a stability criterion in the GALFORM model, shown for this particular realization (solid line) and for a realization where the collapse of 'unstable' discs has not been enforced (dotted line).

It deems discs to be unstable to bar formation if

$$\frac{GM_{\rm disc}/r_{\rm D}}{v_{\rm c}^2} > 1.56. \tag{29}$$

When this inequality is met, all disc material is transferred to a central bulge.

In the usual GALFORM model, some set fraction of disc mass is also transferred to a central black hole (Bower et al. 2006). This was used to arrive at the value in (29) which was found to best match the Magorrian relation between bulge mass and black hole mass, $M_{\rm BH} \sim M_{\rm bulge}^{1.12}$, observed by Häring & Rix (2004). In this study, the black hole accretion fraction is set to zero for the purposes of consistency with the simulation, where this aspect of evolution is not considered. Being physically well motivated, the condition (29) has been applied throughout this study. Conveniently, the precise formalism used need not cause too much concern here; the change that its application produces in the component masses (in Fig. 11, for example) is barely enough to be noticed. This can be understood from Fig. 9, which shows that only a very limited adjustment to the disc mass is required to prevent the limit in (29) from being reached. The mass aggregation of this particular system, and the fact that the specific angular momentum of the gas is relatively high[11] ($\lambda > 0.07$ for the latter half of the halo's history) leads to the stabilization of the system, even if the redistribution of mass is not enforced.

4 RESULTS

4.1 Matched model

4.1.1 Structure

The size and rotation speed are shown for both realizations in Fig. 10. The predicted GALFORM value is based on conserving the initial angular momentum (18) of in-falling gas; an assumption which is particularly effective in this case where the distributions for the initial and final mass were appropriate (Fig. 4). This correspondence is encouraging for the approach of calculating disc sizes from the spin parameter, λ, of the halo from which the gas cooled.

---

[11] The criterion (29) can also be written as

$$M_{\rm disc} > 1.2[M_{\rm DM}(r_{1/2}) + M_{\rm sph}(r_{1/2})], \tag{30}$$

where $M_{\rm DM}$ and $M_{\rm sph}$ are the masses of the DM and spheroid components. The disc's eventual half-mass radius will enclose more DM if the original spin parameter is higher.
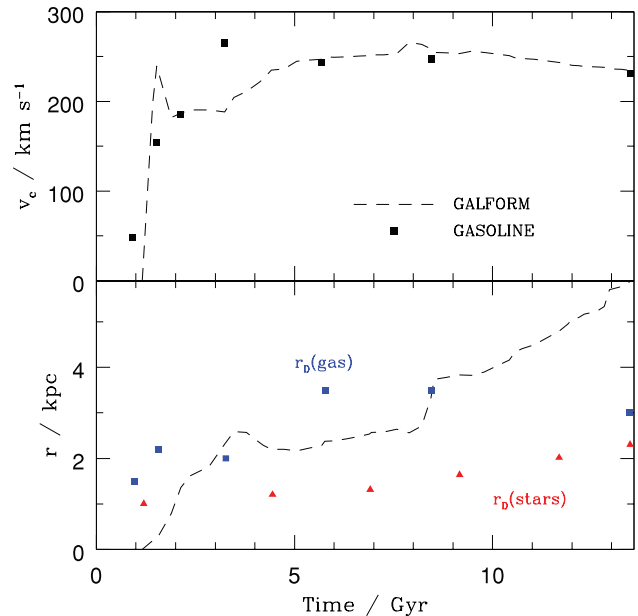

Figure 10. The evolution of the disc rotation velocity, $v_{\rm c}$, (top panel) and the disc scalelength, $r_{\rm D}$, (lower panel) for both realizations. The circular velocity of the simulated disc (black squares) is taken at five comoving kpc from the centre. The approximate gas scalelengths (blue squares) are found by fitting (20) to the simulated gas distributions in Fig. 4. The stellar scalelengths (red triangles) are taken from fitting (20) the stellar profiles, shown later in Fig. B1. Dashed lines are the values calculated in GALFORM from conservation of angular momentum (Section 3.6 – gas and stars are assumed to have the same scalelength).
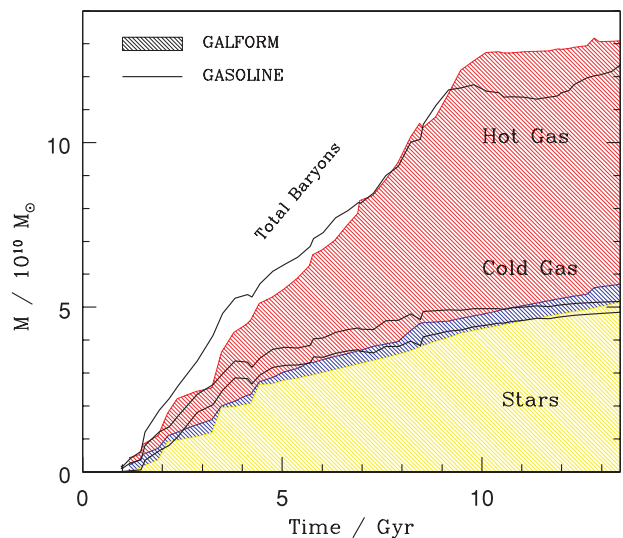

Figure 11. The evolution of hot halo gas, cold disc gas and stars according to both modelling approaches. The star formation rate in GALFORM is calculated using assumptions identical to those in the simulation (21a), but taking into account their integrated effect over the whole disc profile (20).

4.1.2 Cooled mass

The total cooled mass (stars and cold gas), calculated using the methods of Section 3.2, can be seen from Fig. 11. Agreement with the simulation is excellent, given the inherent difference in the methods of calculation, and bolsters confidence in the rather

intricate calculation of the energy radiated by the system through-out its complex merger history (12–17).

The division of the cooled mass into stars and gas are also shown in Fig. 11. The evident agreement can be understood by reference to Figs 4, 7 and 8, which demonstrate that the considerable complexities of the simulation reduce to relatively simple relationships when integrated over the entire disc.

4.1.3 Discrepancies

There is a minor discrepancy at early times ($t < 8$ Gyr) in the cooled mass generated by the two modelling techniques. This is unlikely to be due to differences in cooling from the hot halo: given the similar density distributions (Fig. 4) and cooling function, the two models should yield similar amounts of cooled gas, all else being equal.

The discrepancy is also not due to a difference in the time of development of a hot gas halo. The top panel of Fig. 4 shows that accreted gas in the GALFORM model transitions from unshocked, cold gas to shocked, hot gas at an age of about 2 Gyr, while the bottom panel shows that this time is also when the GASOLINE model begins to develop a hot gas halo.

This demonstrates that the adopted analytic separation of un-shocked and shocked gas (free-fall limited regime versus cooling limited regime) is an excellent analogue to the transition between cold gas accretion and shocked gas accretion seen in simulations as a galaxy crosses the threshold mass capable of developing a hot halo. It is important to note that this transition has always been included in analytic models of galaxy formation, and thus cold gas accretion at low galaxy mass does not alter the standard picture of galaxy formation.

The discrepancy in cooled mass arises after the development of this hot halo. The simulated galaxy continues to accrete cold gas via filaments that penetrate within the hot gas halo (see Fig. 3 here, fig. 5 of Brooks et al. 2009 and also Ocvirk, Pichon & Teyssier 2008; Dekel et al. 2009; Agertz, Teyssier & Moore 2009), while cold gas accretion ends in the analytic model.

4.2 Adapting GASOLINE to emulate semi-analytics

Brooks et al. (2009) demonstrated that the inclusion of cold gas accretion along filaments leads to an earlier phase of star formation than predicted if all gas was initially shock heated, due to the shorter cooling times on to the central galaxy of the cold gas. Fig. 12 shows that the simulated galaxy does indeed form stars earlier than the analytic model, though the overall discrepancy is never more than a factor of 2.

In an attempt to adapt the *simulations* to include similar physics as the *analytic* model, Fig. 12 shows the effect on the stellar growth in the simulations if all of the cold gas had instead been shocked. The hot gas has cooling times several Gyr longer than the cold gas. This is quantified, and a delay has been added to the formation time of the stars spawned from the cold accreted gas (Brooks et al. 2009). As seen in Fig. 12, this delay leads to a slower build up of stellar mass in the simulations. Comparison with Fig.11 shows this to be in even better agreement with the stellar growth of the analytic model, which neglects this cold gas accretion along filaments.

4.3 Comparison with standard GALFORM assumptions

In order to represent the same physical assumptions as this particular simulation, the semi-analytic code had to be significantly modified
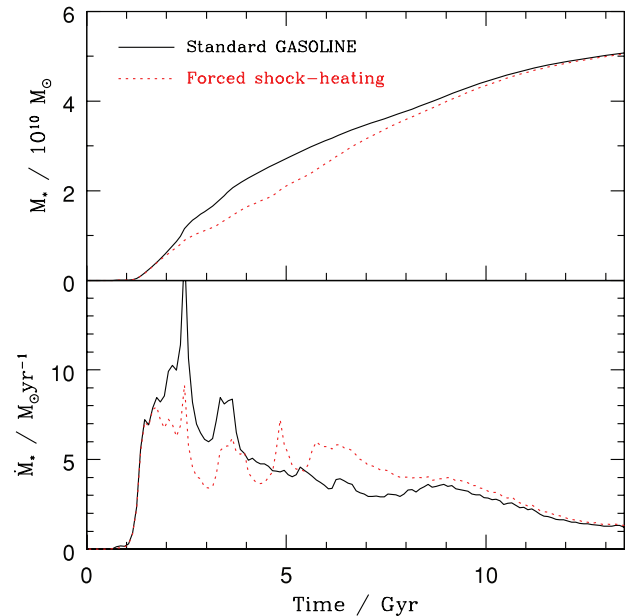

Figure 12. Having focused on modifications to GALFORM, this figure shows the effects of modifying the SPH code GASOLINE to emulate the semi-analytic technique. The solid black line is the star formation rate (lower panel) and its integral (upper panel) which occurs in the standard GASOLINE simulation which has been under discussion. The dotted red lines show the consequence of delaying star formation from cold accreted gas is delayed by a cooling time.

from the version which has been used most recently to study the collective properties of galaxy samples (Bower et al. 2006; Stringer et al. 2009). There are three very significant differences between this version of GALFORM and the simulation studied here, all of which are manifest in the comparison of the history of the simulation and the Bower et al. (2006) model (Fig. 13).

4.3.1 Inclusion of AGN feedback

When there are instabilities (Section 3.9) or merger events in this model which trigger disc collapse, 5 per cent of the available disc gas is assumed to accrete on to a central black hole. If the following criteria are satisfied, it is assumed that no hot halo gas will be able to cool on to the disc.

(i) The resulting Eddington luminosity of the black hole is at least 4 per cent of the cooling luminosity of the halo.

(ii) The cooling time (4) exceeds the free-fall time (5) by some chosen factor (in this case, $t_{cool} > 0.58 t_{ff}$).

This effect is visible in Fig. 13 as a period of rapidly decreasing cold gas mass; during such phases, it is no longer being replenished by gas cooling in from the halo.

4.3.2 Stronger supernova feedback

This is assumed to be extremely strong in the Bower et al. (2006) model, the justification being that a more modest conversion of supernova energy to gas outflow would allow the formation of too many low-mass galaxies. [Indeed, it appears that the predicted number may still be too high (Stringer et al. 2009) even with this generous allocation of feedback energy.]
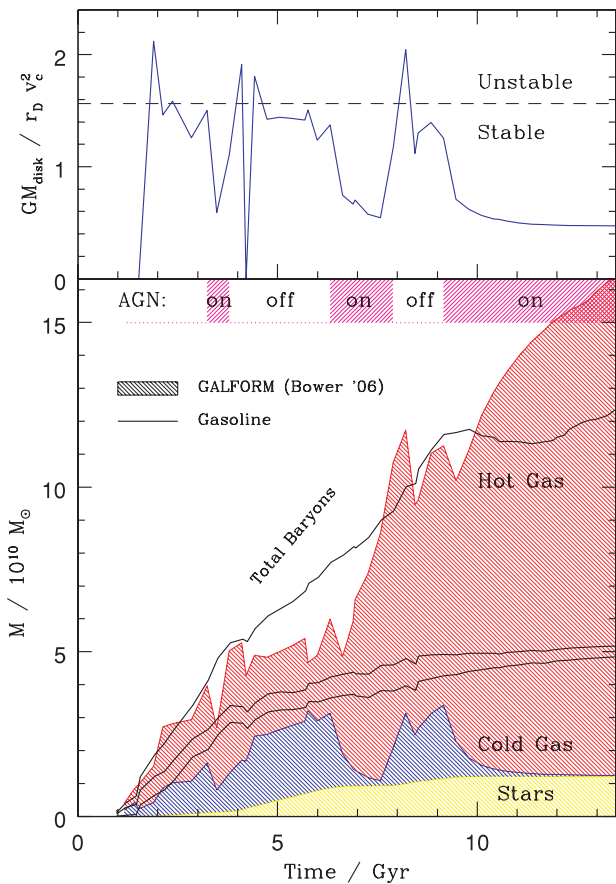
Figure 13. The results of applying the unadjusted version of GALFORM (Bower et al. 2006) to the same merger tree as used for Fig. 11 and indeed the whole paper. To bring out the contrast between the two models, the shading now represents the GALFORM predictions and the lines show the simulation results. The upper panel is the equivalent to Fig. 9, showing the stability of the disc in this version for this model.

In conjunction with the inclusion of AGN heating, this assumption leads to a system with a much greater hot gas component: as soon as the supply of gas to the disc can no longer be replenished by cooling, all the gas in the system quickly ends up in the hot component. This effect is additionally enforced by the low star formation efficiency (see below), which prevents disc mass being locked up into stars.

4.3.3 Inefficient star formation

The star formation assumptions in GASOLINE and the Cole et al. (2000) model have already been contrasted in Section 3.7. The assumed efficiency in the Bower et al. (2006) model is to be so low ($\epsilon_\star = 0.0029$) that, even after allowing for the structure and gas fraction, the final conversion rate is about a factor of 10 *lower* than in the simulation ($\dot{M}_\star/M_{cold} \approx 0$ to $0.2\,\mathrm{Gyr}^{-1}$ as opposed to the rate of 0.5 to $3\,\mathrm{Gyr}^{-1}$ seen in Fig. 7).

This stands out in Fig. 13. The large mass of cold gas which is supported by the disc in the semi-analytic realization (except when cooling is shut off by AGN) cannot be sustained in the simulation due to much more effective conversion to stars.

5 SUMMARY

The merger history of a simulated, Milky-Way-like galaxy has been used to recreate the same system using the GALFORM semi-analytic model.

The incoming trajectories of satellite haloes were investigated and found to be consistent with the usual, assumed distributions. Information on the different modes of accretion on to the main halo is shown to be inherently contained in the merger tree (and would be similarly contained in a statistically generated merger tree which may be used in the study of larger volumes).

The subsequent fate of accreted gas is not modelled equivalently by the two methods, but the analytic framework *does* account for both shocked and unshocked gas. Furthermore, the estimated formation time of the hot halo is in excellent agreement with the simulation.

The single-parameter analytic distributions which are assumed in GALFORM are found to be a satisfactory description of the radial distribution of gas and stars inside the simulated halo. It was also shown that the choice of hot halo scalelength is of minor importance in the calculation of cooled gas mass.

However, the enforcement of spherical symmetry in these analytic distributions means that any subsequent cold accretion along filaments (persisting in spite of shocks elsewhere) is neglected. Thus, the *total* quantity of unshocked gas on to the central galaxy is underestimated, as expected by Brooks et al. (2009). Some gas is consequently delayed in its arrival at the central galaxy, though it is found to proceed at later times.

The disc structure is predicted analytically by conserving the angular momentum of the cooled gas. Resulting scalelengths are within a factor of 2 of the simulation results and the circular velocities even closer. This is quite satisfactory for the case of a single halo, but more work is needed to establish whether the two techniques will consistently agree to this level for an arbitrary set of initial conditions (and whether any disagreement is biased).

With regard to the subsequent evolution of the disc, the GALFORM code was altered to apply the same fundamental assumptions as the simulation by considering their *integrated effect over the assumed disc structure*. Subsequent agreement was excellent given this immense reduction in complexity from many thousands of SPH particles to a few one-dimensional equations.

So, by assuming the same physics and using the same initial conditions, the two techniques predict a final system which is recognizably the same, despite fundamental differences in the way that these assumptions are applied and the evolution followed. This suggests that equations which attempt to characterize the emergent behaviour of the system may, with sufficient understanding, become as reliable as those which emulate the properties of particles themselves.

With such *potential* consistency thus established, the same system was then evolved under physical assumptions which have been more frequently adopted within GALFORM (Bower et al. 2006). The resulting system (at any time) is *not* recognizably the same as that predicted by the simulation: having less than a quarter of the stellar mass, barely half the rotational velocity and with negligible star formation at $z = 0$. Though not all aspects of a system need to be correctly represented to learn from the modelling exercise, this divergence from the same initial conditions proves that at least one of these published models poorly represents the true nature of galaxy formation.

It is hoped that this detailed comparison between two very different theoretical approaches will promote the awareness of their

respective limitations and help lead them both towards a truer description of real galaxies.

REFERENCES

Agertz O., Teyssier R., Moore B., 2009, MNRAS, 397, 64
Benson A. J., 2005, MNRAS, 358, 551
Benson A. J., Bower R., 2010, MNRAS, 405, 1573
Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001, MNRAS, 320, 261
Birnboim Y., Dekel A., 2003, MNRAS, 345, 349
Bower R. G., Coles P., Frenk C. S., White S. D. M., 1993, ApJ, 405, 403
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Brooks A. M., Governato F., Booth C. M., Willman B., Gardner J. P., Wadsley J., Stinson G., Quinn T., 2007, ApJ, 655, L17
Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396
Cattaneo A. et al., 2007, MNRAS, 377, 63
Cole S., Lacey C., Baugh C., Frenk C., 2000, MNRAS, 319, 168
Croton D. J. et al., 2006, MNRAS, 365, 11
Dekel A. et al., 2009, Nat, 457, 451
Efstathiou G., Lake G., Negroponte J., 1982, MNRAS, 199, 1069
Gill S. P. D., Knebe A., Gibson B. K., 2004, MNRAS, 351, 399
Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, MNRAS, 374, 1479
Governato F., Mayer L., Brook C., 2008, in Funes J. G., Corsini E. M., eds, ASP Conf. Ser. Vol. 396, Formation and Evolution of Galaxy Disks. Astron. Soc. Pac., San Francisco, p. 453
Governato F. et al., 2009, MNRAS, 398, 312
Gross M. A. K., 1997, PhD thesis, Univ. California, Santa Cruz
Haardt F., Madau P., 1996, ApJ, 461, 20
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Hatton S., Devriendt J. E. G., Ninin S., Bouchet F. R., Guiderdoni B., Vibert D., 2003, MNRAS, 343, 75
Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., 2003, MNRAS, 338, 903
Jiang C. Y., Jing Y. P., Faltenbacher A., Lin W. P., Li C., 2008, ApJ, 675, 1095
Katz N., White S. D. M., 1993, ApJ, 412, 455
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188
Kennicutt R. C., 1998, Annu. Rev. Astron. Astrophys., 36, 189
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
Kroupa P., Tout C. A., Gilmore G., 1993, MNRAS, 262, 545
Lacey C., Cole S., 1993, MNRAS, 262, 627

Maiolino R. et al., 2008, A&A, 488, 463
Mayer L., Governato F., Kaufmann T., 2008, Adv. Sci. Lett., 1, 7
Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Pontzen A. et al., 2008, MNRAS, 390, 1349
Press W. H., Schechter P., 1974, ApJ, 187, 425
Quinn T., Katz N., Efstathiou G., 1996, MNRAS, 278, L49
Read J. I., Mayer L., Brooks A. M., Governato F., Lake G., 2009, MNRAS, 397, 44
Sharma S., Steinmetz M., 2005, ApJ, 628, 21
Springel V., 2005, MNRAS, 364, 1105
Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074
Stringer M. J., Benson A. J., 2007, MNRAS, 382, 641
Stringer M. J., Benson A. J., Bundy K., Ellis R. S., Quetin E. L., 2009, MNRAS, 393, 1127
Sutherland R. S., Dopita M. A., 1993, ApJ, 88, 253
Tormen G., 1997, MNRAS, 290, 411
Wadsley J. W., Stadel J., Quinn T., 2004, New Astron., 9, 137
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Yoshida N., Stoehr F., Springel V., White S. D. M., 2002, MNRAS, 335, 762
Zolotov A., Willman B., Brooks A. M., Governato F., Brook C. B., Hogg D. W., Quinn T., Stinson G., 2009, ApJ, 702, 1058

APPENDIX A: SATELLITES

The orbits of subhaloes around their host can be characterized (Lacey & Cole 1993) by the dimensionless parameter

$$\Theta = \left(\frac{j}{j_c(E)} \right)^{0.78} \left(\frac{r_c(E)}{r_v} \right)^2 , \tag{A1}$$

where $r_c(E)$ and $j_c(E)$ are the radius and specific angular momentum of a circular orbit with energy E. For a potential with a flat rotation curve; where the circular velocity at all radii is v_c,

$$\Phi(r) = \Phi(r_v) + v_c^2 \ln\left(\frac{r}{r_v} \right) . \tag{A2}$$

Hence, for a subhalo at position, $\boldsymbol{r}$ with velocity $\boldsymbol{v}$, the radius of a circular orbit with the same energy is

$$r_c(\boldsymbol{v}, \boldsymbol{r}) = |\boldsymbol{r}| \exp\left[\frac{\boldsymbol{v}^2}{2 v_c^2} - \frac{1}{2} \right] . \tag{A3}$$

The orbital parameters are found by simply applying equations (A1) to (A3) to the position and velocity vectors of each satellite halo (after moving to a frame where these vectors are both **0** for the host halo):

$$\Theta = \left(\frac{j}{v_c r_v} \right)^{0.78} \left(\frac{r_c(E)}{r_v} \right)^{1.22}$$
$$= \left(\frac{|\boldsymbol{v} \times \boldsymbol{r}|}{v_c r_v} \right)^{0.78} \left(\frac{|\boldsymbol{r}|}{r_v} \right)^{1.22} \exp\left[0.61 \left(\frac{\boldsymbol{v}^2}{v_c^2} - 1 \right) \right] . \tag{A4}$$

The GALFORM model assigns each satellite a value for this parameter, Θ, drawn at random from the log-normal distribution that it was found to follow in simulations by Tormen (1997):

$$\frac{\mathrm{d} \ln n}{\mathrm{d} \ln \Theta} = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{(\ln \Theta + 0.14)^2}{2\sigma^2} \right] , \tag{A5}$$

with $\sigma = 0.26$. The sample of orbit parameters produced by this random assignment is shown in Fig. A1, together with the actual values from the Brooks et al. (2009) simulation.
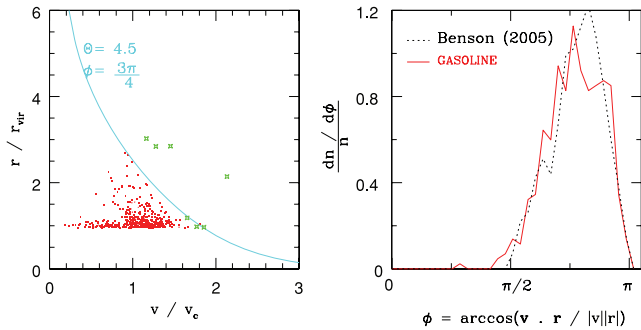
Figure A1. Left-hand panel: the velocities and radii of the satellite haloes in the simulation of Brooks et al. (2009), at the time-step immediately before they are deemed to have merged with their host. The dots are those which appear in the distribution of Θ in Fig. 1, crosses are those that have $\Theta > 4.5$ and hence do not appear. The line shows the constant value of Θ corresponding to that cut-off, for the given approach angle. Right-hand panel: the distribution of approach angles of these satellites alongside the distribution found from the study by Benson (2005).

The fraction of the halo free-fall time that it takes for the satellite to merge through dynamical friction is assumed to be proportional to Θ:

$$\frac{\tau_{mrg}}{[G\rho_v(z)]^{-\frac{1}{2}}} = \Theta \frac{0.3722}{\ln(M_{host}/M_{sat})} \frac{M_{host}}{M_{sat}}. \tag{A6}$$

This is based on the standard Chandrasekhar formula for the dynamical friction and appears originally in Lacey & Cole (1993). ρ_v is the mean density at the virial radius (determined by the cosmology, not the specific halo's properties).

The distribution of angles at which substructures enter their host halo has been studied by Benson (2005) for the VLS and VIRGO simulations and found to have a repeatable distribution. This distribution can be applied in the GALFORM model to generate an alternative set of orbital parameters, Θ, which are shown in Fig. A1 alongside the standard assumption of the log-normal distribution, and the results from GASOLINE.

Merger times predicted by (A6) have been compared by Jiang et al. (2008) with the results of GADGET2 simulations (Springel 2005). The agreement found was of the order of a factor of 2, which is deemed acceptable for the continued use of this formula in GAL-FORM. However, since (A6) was found to consistently underestimate the simulated merger time, the improved fitting formula proposed by Jiang et al. (2008) may well be adopted in future.

Unfortunately, due to the very different definition of merging adopted by the halo finder used here (Section 3.1), a meaningful comparison of the respective times in the two realization has not been possible.

APPENDIX B: DISC PROFILES

The stellar disc radii that appear in Fig. 5 are generated from analysis of the distribution of stars in the simulation. Fig. B1 compares the
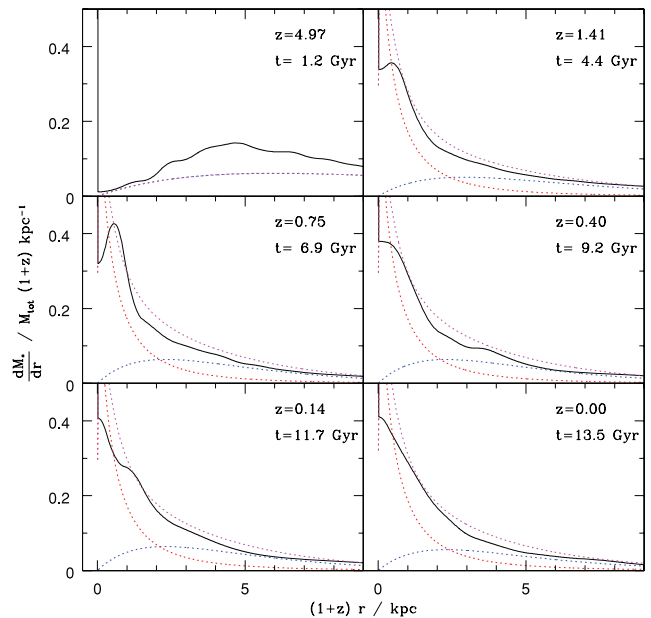

Figure B1. The stellar mass profile from the simulation of Brooks et al. (2009) at six different redshifts (solid black lines). The dotted blue lines show the analytic forms which are assumed by the GALFORM model for the disc (20). Dotted red lines show the bulge components (B1). The total stellar contribution, shown by dotted magenta lines, is constrained to integrate to the same mass as the solid lines. The two lines are therefore fitted using three free parameters: The relative fraction of mass in the two components, the disc scalelength, r_D and the scalelength of the bulge profile, r_b.

distribution of stellar mass from the simulation with the analytic forms assumed by GALFORM, both for disc stars (20) and for the mass of stars in the bulge, assumed to be distributed such that the projected surface density profile is given by

$$\Sigma_{bulge} \propto \exp\left[-\left(\frac{r}{r_b}\right)^{\frac{1}{4}}\right]. \tag{B1}$$

Though the agreement is not comprehensive, the analytic forms are an adequate description of the global distribution of simulated stars for the majority of the systems evolution. Unsurprisingly, the simple profiles of (20) and (B1) fail to describe the simulated system at early times, before an ordered galactic system has formed.

The fact that GALFORM adheres, at *all* times, to the fitting forms that are really only appropriate to recent times ($z < 2$) is clearly a valid criticism of this technique. However, as long as such forms are indeed a good description of low-redshift systems, then the only way their earlier misjudgements can significantly mislead the predictions of the model is in the early time star formation rates (Fig. 7), which are based on this assumed structure.