Mon. Not. R. Astron. Soc. **343,** 679–691 (2003)

The effects of photoionization on galaxy formation – III. Environmental dependence in the luminosity function

A. J. Benson,[1⋆] C. S. Frenk,[2] C. M. Baugh,[2] S. Cole[2] and C. G. Lacey[2]

[1] *California Institute of Technology, MC 105-24, Pasadena, CA 91125, USA*
[2] *Physics Department, University of Durham, Durham DH1 3LE*

Accepted 2003 April 11. Received 2003 April 9; in original form 2002 October 16

ABSTRACT

Using semi-analytic modelling techniques, we calculate the luminosity function of galaxy populations residing in cold dark matter haloes of different masses. We pay particular attention to the influence of the re-ionization of the Universe on the number of faint galaxies, and to the effects of dynamical friction and tidal limitation of satellites on the number of bright galaxies. We find substantial differences in the shapes of the galaxy luminosity functions in haloes of different mass, which reflect generic features of the cold dark matter model of galaxy formation, and thus offer the opportunity to test it. We then consider how the individual halo luminosity functions combine to produce the global luminosity function. Surprisingly, the global function ends up having a shallower faint-end slope than those of the constituent halo luminosity functions. We compare our model predictions with the limited observational data sets compiled by Trentham & Hodgkin. We find good agreement with the luminosity functions measured in the Virgo and Coma clusters, but find significant disagreement with the luminosity functions measured in the Local Group and Ursa Minor cluster. We speculate on possible inadequacies in our modelling and in the existing observational samples. The luminosity functions of galaxies in groups and clusters that have been identified in the Two Degree Field (2dF) and Sloan Digital Sky Survey (SDSS) galaxy redshift surveys offer the prospect of testing galaxy formation models in detail.

Key words: galaxies: formation – galaxies: luminosity function, mass function – cosmology: theory.

1 INTRODUCTION

Understanding the galaxy luminosity function has been a goal of galaxy formation theory for several decades (e.g. White & Rees 1978; Cole 1991; White & Frenk 1991). A particularly interesting question concerns whether the luminosity function is universal, or whether it depends on environmental factors, such as the mass of the dark halo that hosts a particular galaxy population. Considerable attention has been paid to the faint end of the luminosity function, which has a much flatter slope than that of the low-mass end of the halo-mass function that has been predicted in cold dark matter (CDM) models of galaxy formation (e.g. Norberg et al. 2002; Benson et al. 2003).

The early work of White & Rees (1978) showed that the number of faint galaxies must have been strongly affected by feedback processes, which prevented most of the gas from cooling in small haloes at early times. Some likely feedback mechanisms – such as the injection of energy into the interstellar medium in the course of stellar evolution – depend on the internal properties of the galaxy, and so their effects may be expected to be independent of the large-scale environment. Indeed, a number of observational studies – such as a recent analysis of the Two Degree Field (2dF) galaxy redshift survey (De Propris et al. 2002) – find no significant difference between the luminosity functions of galaxies in rich clusters and in the field. Other studies, however, have found the opposite. For example, Phillipps & Shanks (1987) concluded that galaxies in rich clusters have luminosity functions with considerably steeper faint ends than galaxies in the field. More recently, Trentham & Hodgkin (2002) have claimed that the faint end of the galaxy luminosity function varies systematically with environment, increasing in slope from small, diffuse systems such as the Local Group, to massive, dense systems such as the Coma cluster.

In the CDM model of galaxy formation, dark matter haloes retain considerable substructure after they collapse and virialize (e.g. Klypin et al. 1999; Moore et al. 1999), and some of these subhaloes are associated with sites of galaxy formation. The mass function of subhaloes appears to be approximately independent of the mass of the parent halo (when scaled appropriately by the mass of the parent halo; Moore et al. 1999). Thus, as first noted by Kauffmann, White & Guiderdoni (1993), trends such as those inferred by Trentham & Hodgkin (2002) would require processes that either preferentially suppress the formation of dwarf galaxies in low-mass

⋆E-mail: abenson@astro.caltech.edu

parent haloes, or destroy them after they form. An effective mechanism for suppressing the formation of small galaxies is the reheating of the intergalactic medium (IGM), caused by the re-ionization of the Universe at a redshift $z \gtrsim 6$. This inhibits the formation of dwarf galaxies after the re-ionization epoch. Tully et al. (2002) have argued that this process could introduce an environmental dependence into the galaxy luminosity function, because the fraction of low-mass progenitor haloes (in which dwarf galaxies form) that are assembled before re-ionization is larger in present-day clusters than in lower-mass systems. As a consequence, clusters might be expected to have a higher ratio of dwarf to giant galaxies than low-mass systems such as groups.

The effect of re-ionization on the formation of galaxies has been the subject of several recent studies [Bullock et al. 2000; Benson et al. 2002 (hereafter Paper I); Somerville 2002], aimed mostly at investigating the discrepancy between the large number of subhaloes found in N-body simulations of galactic CDM haloes and the small number of satellite galaxies observed in the Local Group. In this paper, we employ the galaxy formation model of Paper I to calculate the luminosity function of galaxy populations residing in CDM haloes of different mass. We find that there are significant differences in these luminosity functions, and we then explore how they combine together to build up the global luminosity function, with particular emphasis on the faint-end slope. Diaferio et al. (1999) carried out a study of luminosity functions in haloes of different mass, using a semi-analytic model of galaxy formation. The present work considerably extends upon that of Diaferio et al. (1999), but there is generally good agreement between the two where they overlap.

To calculate galaxy luminosity functions in haloes of different mass correctly, it is important to include tidal effects on satellite galaxies, as these are a potential galaxy-destruction mechanism. Our model treats these effects in considerably more detail than previous models of galaxy formation have done. We find that tidal effects are important in limiting the formation of massive galaxies at the centres of rich clusters.

In this paper, we compare the results of our calculations to the data of Trentham & Hodgkin (2002), and assess whether feedback from re-ionization is a viable explanation of the trend claimed by these authors. The existing data set is small, but forthcoming results from the 2dF and SDSS surveys will enable much more extensive comparisons with the theory.

The remainder of this paper is arranged as follows. In Section 2 we briefly outline our model of galaxy formation; in Section 3, we present results for the environmental dependence of the luminosity function; and in Section 4, we compare our model with the available observational data. Finally, in Section 5, we present our conclusions. In the Appendices, we present several simple models of photoionization suppression, to elucidate how this mechanism works.

2 MODEL

We employ the semi-analytic model of galaxy formation, described in detail by Cole et al. (2000) and in Paper I, to compute the properties of galaxies in a range of environments at $z = 0$. The reader is referred to those papers for a complete and detailed description of the model. Briefly, the hierarchical formation of dark matter haloes is calculated using the extended Press–Schechter formalism (Press & Schechter 1974; Bower 1991; Bond et al. 1991). The formation of galaxies in the resulting dark matter halo merger trees is followed by means of simple, physically motivated models of gas cooling, star formation and galaxy merging. Recent work has demonstrated that, at least in so far as gas cooling is concerned, these simplified calculations are in excellent agreement with the results of N-body/hydrodynamical simulations (Benson et al. 2001a; Yoshida et al. 2002; Helly et al. 2003). Applying a stellar population synthesis model gives galaxy luminosities in different passbands. The model includes a prescription for the feedback from supernovae, which drives gas out of galaxies at a rate proportional to the current star formation rate, with a constant of proportionality that is larger for less strongly-bound systems. This negative feedback flattens the faint-end slope of the luminosity function, but, with the parametrization adopted by Cole et al. (2000), its effect is not strong enough to account for the measured slope (Paper I). Of course, this model for the feedback from supernovae is purely phenomenological, and so it remains possible that a physically correct model of supernovae feedback would be able to produce a sufficiently flat faint-end slope.

Paper I developed – and incorporated into the Cole et al. (2000) model – a detailed, self-consistent treatment of photoionization suppression of galaxy formation (hereafter PhS). By reducing the number of faint galaxies further, this mechanism brought the model luminosity function into excellent agreement with observations. In Paper I, we also included a much more detailed treatment (than was previously possible) of the evolution of satellite galaxies experiencing dynamical friction and mass loss due to tidal forces as they orbit within larger haloes. (Hereafter, we abbreviate 'tidal limitation' as TiL.) This also affects the shape of the luminosity function, and so we will explore this process in this paper.

We adopt essentially the same model parameters as in Paper I, but with the following differences. Whereas the models of Paper I used the Press–Schechter halo mass function, we choose to adopt the function proposed by Jenkins et al. (2001), which gives a better match to the results of N-body simulations.[1] Consequently, to produce a reasonable galaxy luminosity function, we find it necessary to adjust the parameters Ω_b from 0.020 to 0.024, $\epsilon_\star$ (which determines the star formation rate) from 0.0050 to 0.0067, and Υ (which affects mass-to-light ratios) from 1.38 to 1.03 (because Υ affects the fraction of mass recycled in star formation, we adjust that fraction accordingly). Even with these changes, the mass function of Jenkins et al. (2001) results in a somewhat worse fit to the bright end of the luminosity function than was obtained in Paper I. We defer further study of this aspect of the model to future work. We adopt an escape fraction of ionizing photons from galaxies, f_esc, of 12 per cent. This produces a neutral hydrogen fraction for gas at zero overdensity in the IGM of 5×10^{-4} at $z = 6$, in agreement with the lower limit of 3.4×10^{-4} derived by Lidz et al. (2002), while causing re-ionization to occur at as high a redshift as possible. [Note that our models are therefore inconsistent with the high redshift of re-ionization implied by the optical depth measurements of Kogut et al. (2003), based on the observations of the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) satellite. We prefer, in this work, to use this lower re-ionization redshift, because the filtering-mass evolution in models with a high re-ionization redshift is as yet unexplored.] This maximizes the associated suppression of galaxy formation.[2] The resulting model is essentially the same as the $f_\mathrm{esc} = 10$ per cent model that was presented in Paper I.

[1] However, we retain the standard extended Press–Schechter scheme, described by Cole et al. (2000), to produce merger trees for each halo.

[2] We could increase f_esc slightly without violating the limit of Lidz et al. (2002), but, because the neutral hydrogen fraction drops very rapidly in our model at these redshifts (Paper I), it makes almost no difference to the suppression of galaxy formation.

3 THE GALAXY LUMINOSITY FUNCTION

3.1 Luminosity functions in haloes of different mass

We begin by studying the luminosity functions of galaxies in individual dark matter haloes of different mass, and then proceed to consider how these combine together to form the global galaxy luminosity function. Using the model described in Section 2, we simulate galaxy formation in a large number of haloes of masses 10^{11}, 10^{12}, 10^{13}, 10^{14} and $10^{15}\,h^{-1}\,\mathrm{M}_\odot$.[3] This choice includes halo masses appropriate to systems ranging from the Local Group to rich galaxy clusters. (The $10^{11}\,h^{-1}\,\mathrm{M}_\odot$ bin corresponds to haloes less massive than that of the Local Group, but we include it because it shows the most extreme effects.) Merger trees are constructed with a resolution of 10^{-5} of the final halo mass or $10^{8}\,h^{-1}\,\mathrm{M}_\odot$ – whichever is smaller. This is sufficient to follow accurately the formation and evolution of galaxies brighter than $M_\mathrm{B} - 5\log h = -10$, except in the most massive clusters ($10^{15}\,h^{-1}\,\mathrm{M}_\odot$), for which the resolution begins to affect our results for $M_\mathrm{B} - 5\log h = -12$ and fainter; for these clusters, we show results only down to $M_\mathrm{B} - 5\log h = -13$. We perform calculations for our 'standard model' (i.e. including the effects of PhS and TiL of satellite galaxies) and repeat them, switching off PhS and/or TiL while keeping all other parameters fixed (resulting in a total of four different models), in order to assess the effects of each physical mechanism separately. We simulate a sufficiently large number of haloes of each mass, so that the average luminosity functions are reasonably smooth.

The solid lines in Fig. 1 show the luminosity functions of galaxies in our simulated haloes. The basic form of the standard model luminosity function is similar in all masses of halo. It has two components: a near-power-law distribution at faint magnitudes, which is made up of satellite galaxies (i.e. those galaxies that do not reside at the centre of the final halo); and a strongly peaked distribution at bright magnitudes, beyond which no more galaxies are found and which is made up of the central galaxy in each final halo. As we consider haloes of increasing mass, both components shift to ever brighter luminosities, but the peaked component decreases in amplitude relative to the power-law component. This indicates that, in low-mass haloes, a single galaxy tends to dominate; in clusters, the light is more equally shared among many galaxies.

The origin of the two components that contribute to the galaxy luminosity function can be traced to the distribution of substructure in the dark matter halo. Fig. 2 shows the distribution of dark matter halo masses (prior to any mass loss due to TiL) associated with galaxies in $10^{12}\,h^{-1}\,\mathrm{M}_\odot$ systems at $z = 0$ (solid histogram). The peak close to $10^{12}\,h^{-1}\,\mathrm{M}_\odot$ is associated with the halo as a whole, and therefore with the central galaxy. The distribution shows the same two components as the luminosity function: a peak associated with the central galaxy and a near-power-law distribution of substructure halo masses. The dashed histogram shows the corresponding mass distribution when merging is switched off (i.e. substructure haloes are assumed to survive forever inside their host halo). Even in this case, the two components remain clearly visible, demonstrating that their origin is intimately linked to the nature of the hierarchical growth of haloes in a CDM cosmology and is not an artefact of our detailed treatment of mergers. Two components are also visible in the substructure mass distributions of more massive haloes, such as clusters. However, in this case, galaxy formation at the centre of the main halo is rather inefficient because of the long cooling time of gas and, as a result, the peak component is not prominent in the luminosity function. Our results are qualitatively similar to those obtained by Kauffmann et al. (1993) for luminosity functions in haloes with the masses of the Milky Way and the Virgo cluster, calculated using a different semi-analytic model.

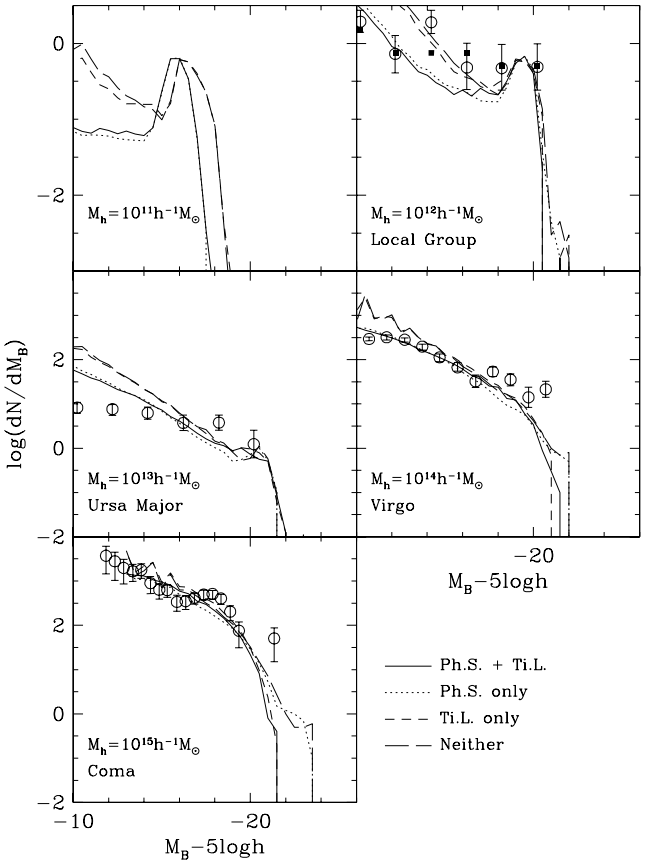

Figure 1. *B*-band galaxy luminosity functions in haloes of different mass. Each panel shows the mean predicted luminosity function in an ensemble of dark matter haloes, with the mass given in the legend (ranging from small galaxies to clusters). Specifically, we show the mean number of galaxies per magnitude, per halo. Solid lines include photoionization suppression (PhS) and tidal limitation (TiL); dotted lines have PhS only; short-dashed lines have TiL only; and long-dashed lines have neither. Open circles show observed luminosity functions from the compilation of Trentham & Hodgkin (2002), as indicated in the legend of each panel. These have been normalized arbitrarily, to permit easier comparison of their shapes with the models. Filled squares in the $10^{12}\,h^{-1}\,\mathrm{M}_\odot$ panel show the luminosity function compiled by Benson et al. (2002b), which includes only galaxies identified as lying within the virial radius of the dark halo of either the Milky Way or M31. The comparison between the model and the observational data is discussed in Section 4.

Observed luminosity functions, $\phi(L) = \mathrm{d}n/\mathrm{d}L$ (where n is the number of galaxies per unit volume and L is luminosity), are typically described by the Schechter function,

$$\phi(L)\,\mathrm{d}L = \phi_* \left(\frac{L}{L_*}\right)^{\alpha} \exp\left(-\frac{L}{L_*}\right) \frac{\mathrm{d}L}{L_*}, \qquad (1)$$

where ϕ_*, L_* and α are parameters. At magnitudes faint compared with the characteristic luminosity, L_*, this function has the simple form $\phi(L) \propto L^{\alpha}$. Following normal practice, we will refer to α as 'the faint-end slope'. Although the luminosity functions in Fig. 1 are not well-fitted by the Schechter form, we can still obtain a formal

[3] Throughout this paper we write the Hubble constant as $H_0 = 100\,h$ km s^{-1} Mpc^{-1}.

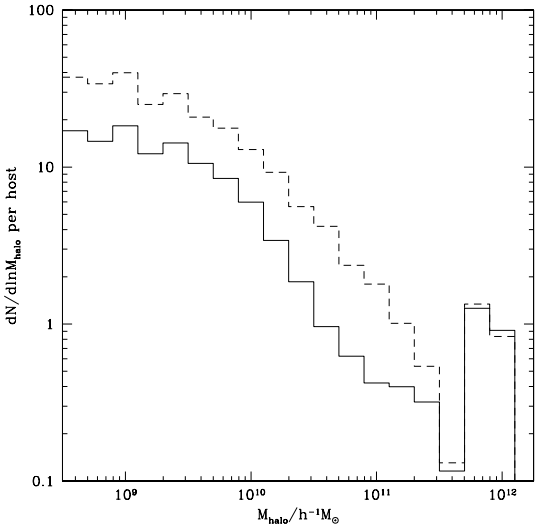

Figure 2. The mass function of haloes associated with galaxies located inside haloes of total mass $10^{12}\,h^{-1}\,\mathrm{M}_\odot$ at $z = 0$. Masses are shown prior to any mass loss induced by TiL. The solid histogram shows the results without the standard merging scheme, whereas the dashed histogram shows the results when no merging is included.

value of α by fitting to the data plotted in the first few bins. We find slopes in the range $-1.57 < \alpha < -1.48$, except in the lowest mass haloes ($10^{11}\,h^{-1}\,\mathrm{M}_\odot$), for which we find a much flatter slope of $\alpha = -1.06$ (although it is possible that a steeper slope would also be obtained in this case, if we probed to even fainter luminosities). There is little evidence for a strong variation of slope with halo mass, except perhaps for a weak (and rather insignificant) trend for *flatter* slopes in higher-mass clusters. However, it is important to note that the slope is not constant with galaxy magnitude. This is clearly seen, for example, in the $10^{12}\,h^{-1}\,\mathrm{M}_\odot$ haloes, for which the slope is close to $\alpha = -1$ in the $-18 < M_\mathrm{B} - 5\log h < -15$ region.

Switching off TiL in our model, and reverting to the more simple calculation of dynamical friction of Cole et al. (2000), results in the dotted curves of Fig. 1. Clearly, this has little influence on the faint end of the luminosity function, resulting in only smaller increases or decreases in the number of galaxies of a given luminosity. A more interesting effect of TiL is apparent in the massive clusters (10^{14} and $10^{15}\,h^{-1}\,\mathrm{M}_\odot$ haloes). Here, switching off TiL results in a new population of very luminous galaxies, visible as a 'bump' in the luminosity functions at bright magnitudes. These highly luminous objects form through multiple mergers at the centres of cluster haloes. When TiL is included, the dynamical friction drag – and thus the merger rate of massive galaxies in clusters – is significantly reduced, and none of these highly luminous objects form. As shown in the left-hand panel of Fig. 3, cluster central galaxies are about an order of magnitude less massive when our accurate calculation of dynamical friction is used. Note that the inclusion of TiL makes much less difference to $10^{12}\,h^{-1}\,\mathrm{M}_\odot$ haloes because of the very different shape of the luminosity function in these systems, in which there is relatively little mass in satellites to merge into the central object. The difference between the merger rates in our model and that of Cole et al. (2000) cannot fully be accounted for simply by increasing the dynamical friction time-scale in the latter model by some factor (even if this factor is a function of halo mass). The shape of the distribution of merger times is significantly different in the two models and, in our calculation, it varies slightly with halo mass.

A similar effect to that revealed by a comparison of the models with and without TiL in Fig. 1 was noticed by Springel et al. (2001) in the luminosity functions of clusters, calculated using a different semi-analytic model. Springel et al. (2001) computed luminosity functions, adopting simple estimates of both galaxy merger rates (similar to those assumed in the dotted curves in Fig. 1) and merger rates, obtained directly from a high resolution N-body simulation of the dark matter component. They found that mergers in the simple scheme produced highly luminous galaxies that were not present in the N-body calculation. Springel et al. traced the cause of this difference to the fact that the simple scheme underestimates the merging time-scale of rather massive subhaloes in clusters (although it predicts the overall merging rate rather well). In Appendix B, we show that the difference between our detailed TiL model and simple merging models can be traced to our choice of Coulomb logarithm in the calculation of dynamical friction forces. Mass loss from haloes, which also reduces the strength of dynamical friction, actually has little effect on the final luminosities of cluster galaxies, as is also demonstrated in Appendix B.

The reduction in the masses of central cluster galaxies in our model may be too strong. With a mass of $\sim 10^{11}\,h^{-1}\,\mathrm{M}_\odot$, the most massive central cluster galaxy that forms is significantly less massive than many cD galaxies (which are typically around $2 \times 10^{11}\,h^{-1}\,\mathrm{M}_\odot$, with examples of up to $10^{12}\,h^{-1}\,\mathrm{M}_\odot$; e.g. Treu & Koopmans 2002, and references therein). Furthermore, as may be seen in the upper right-hand panel of Fig. 4, the luminosity of the central cluster galaxy actually decreases with increasing cluster mass. As a result, the merging model used here does not reproduce the trend of increasing galaxy correlation length with luminosity for the brightest galaxies, seen in the simpler merging model of Cole et al. (2000) (see Benson et al. 2001b). Whereas our model of satellite merging reproduces many aspects of N-body simulations accurately (Benson et al., in preparation), it is clear that our understanding of the formation of central cluster galaxies remains incomplete.

The right-hand panel of Fig. 3 shows the mass functions of *surviving* substructure haloes in clusters at $z = 0$. The dotted histogram corresponds to the model with no TiL, whereas the dashed histogram corresponds to the model with TiL, but the mass of each substructure is plotted as it was *before* it experienced any tidal mass loss. Clearly, the inclusion of TiL has reduced the number of mergers, because there are more surviving substructures of a given initial mass. The solid histogram gives the distribution of substructure masses after tidal mass loss. Comparing the dashed and solid histograms indicates that haloes typically lose between 50 and 70 per cent of their mass to tidal stripping (with higher-mass haloes losing proportionally less than lower-mass haloes).

Switching off PhS (short-dashed lines in Fig. 1) has a more dramatic effect on the luminosity functions than switching off TiL. First, the faint-end slope steepens in almost all systems to values in the narrow range of $-1.71 < \alpha < -1.64$. (The exception are the lowest-mass haloes, which have $\alpha = -1.38$ because of the particularly strong effects of supernovae feedback in these haloes.) The increase in the number of faint galaxies is more pronounced in lower-mass systems. In high-mass systems, there is a weak trend for the bright end of the luminosity function to be shifted slightly faintwards when PhS is ignored. In this case, more gas is locked up in the fainter galaxies, leaving less to form massive, bright galaxies. In the lowest-mass systems, the opposite effect occurs, and the central galaxies become brighter when PhS is ignored. In these $10^{11}\,h^{-1}\,\mathrm{M}_\odot$ haloes, PhS suppresses the amount of gas that is able to accrete into the halo, thus reducing the luminosity of the associated

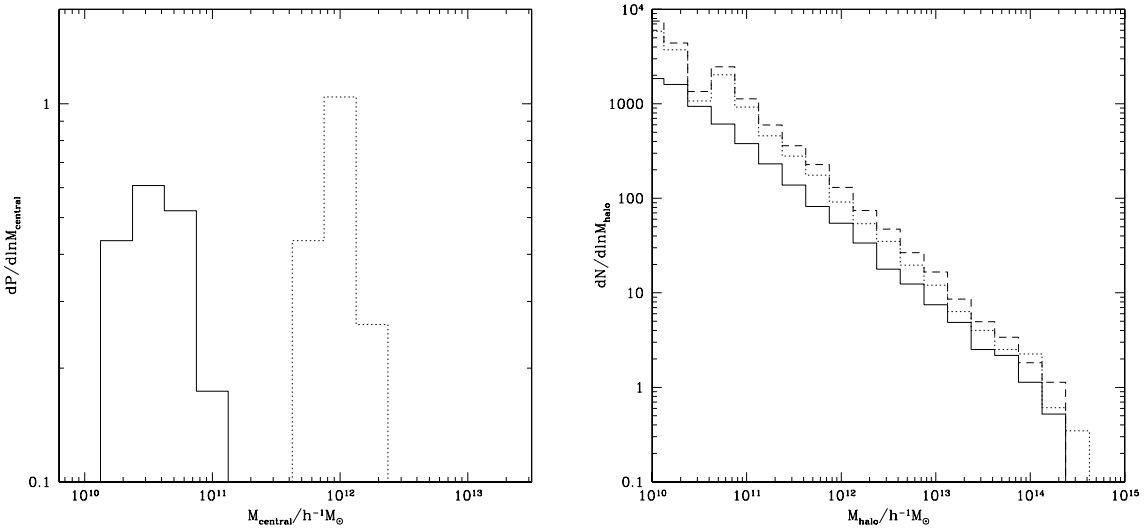

Figure 3. Left-hand panel: the distribution of baryonic masses of galaxies residing at the centres of $10^{15}\,h^{-1}\,\mathrm{M}_\odot$ clusters at $z = 0$. The solid histogram shows results from our standard model which includes TiL, whereas the dotted histogram corresponds to a model with no TiL. Right-hand panel: the mass function of *surviving* substructure haloes in $10^{15}\,h^{-1}\,\mathrm{M}_\odot$ clusters at $z = 0$. The dotted histogram is from a model with no TiL; the dashed and solid histograms are from the model with TiL, and give the masses of haloes before and after tidal mass loss, respectively.

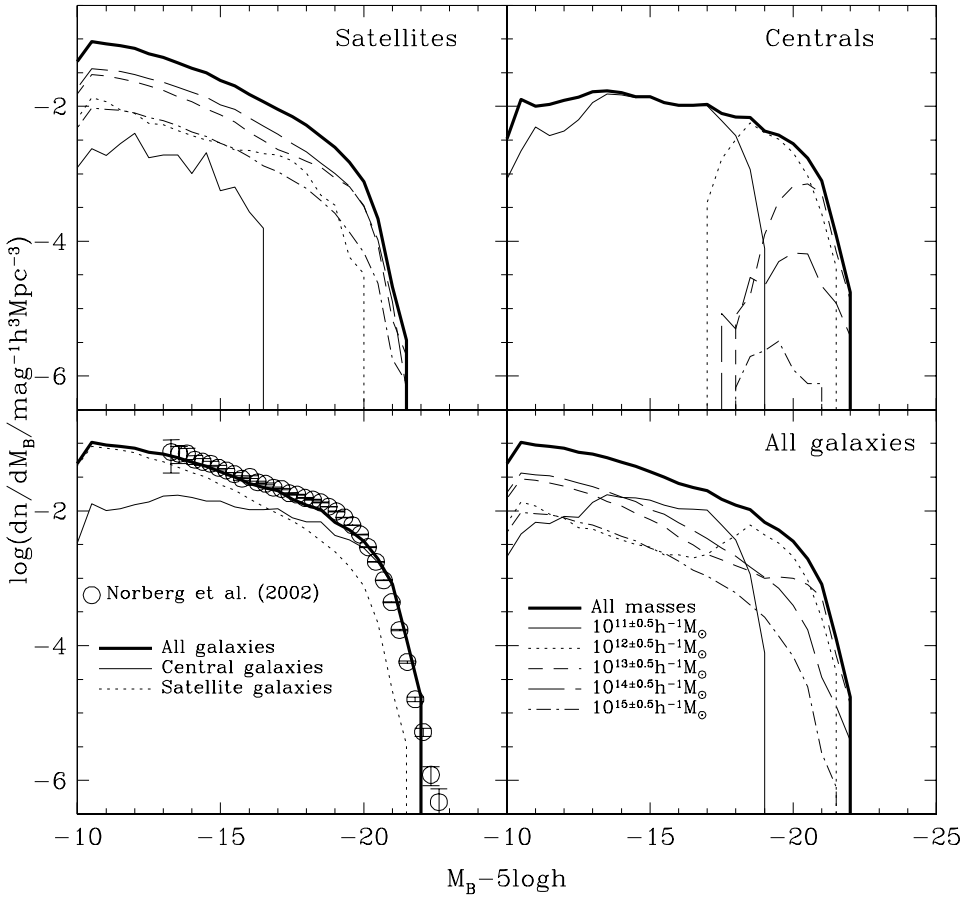

Figure 4. *B*-band galaxy luminosity functions. Upper left-hand panel: the luminosity functions of satellite galaxies (i.e. those which are not the central galaxy of their host halo). The heavy solid line shows the luminosity functions summed over all halo masses, whereas thin lines show the contributions from haloes in different mass ranges, as indicated by the key in the lower right-hand panel. Upper right-hand panel: the luminosity functions of central galaxies. Line types are as in the upper left-hand panel. Lower left-hand panel: a comparison of the satellite (dotted line) and central galaxy (thin solid line) luminosity functions. Also shown is the total luminosity function (heavy solid line) and the observational determination of the total luminosity function from Norberg et al. (2002). Lower right-hand panel: the contributions from all types of galaxy (i.e. satellites and centrals) in haloes of different mass to the total luminosity function. Line types are as in the upper panels.

galaxy. Kauffmann et al. (1993) considered the possibility that the suppression of cooling in haloes with virial circular velocity $V_c < 150$ km s^{-1} could suppress the formation of dwarf galaxies around the Milky Way, and concluded that this was required to explain the data. These results first suggested that PhS may be relevant for the abundance of Local Group dwarfs, although the value of 150 km s^{-1} adopted by Kauffmann et al. (1993) is significantly higher than typically expected from PhS calculations. We have shown here that PhS does indeed suppress the number of dwarf satellites, using our more sophisticated treatment of galaxy merging and assuming modern values for the cosmological parameters. The ways in which various processes associated with photoionization shape the luminosity function in haloes of different mass – particularly its faint end – are discussed in detail in Appendix A. There, we also discuss the effect of various approximations in the treatment of photoionization.

The long-dashed lines in Fig. 1 show the results when both TiL and PhS are ignored. The net effect is essentially an accumulation of the effects seen when either process was switched off individually, and results in a further steepening of the faint-end slopes. In this model, only feedback from supernovae affects the faint end of the luminosity function, in a manner which is essentially independent of the final parent halo mass.

3.2 The global luminosity function

We now explore how the individual halo luminosity functions presented in Section 3.1 combine to produce the global luminosity function. An initially surprising aspect of the individual halo luminosity functions is that their faint-end slopes are, in all cases, steeper than that of the global luminosity function. The latter, as we showed in Paper I, has a slope of around -1.2, both in our models and in the real Universe. How does PhS contrive to produce a flatter faint-end slope for the total population than for any of the constituent individual haloes? To understand this behaviour, we must consider the contribution of haloes of different mass to the global luminosity function.

The upper panels of Fig. 4 show the contributions to the global luminosity function from satellite and central galaxies (left- and right-hand panels, respectively) residing in haloes of different mass. The slope of the combined satellite luminosity function reflects the power-law slope of the individual halo luminosity functions, and is dominated over a wide range of luminosities by haloes of 10^{13}–$10^{14}\,h^{-1}\,\mathrm{M}_\odot$. The luminosity function of central galaxies, on the other hand, has a much flatter faint-end slope, which is determined by the way in which central galaxy luminosity scales with halo mass. In our models, this scaling depends primarily on the combined effects of supernovae and PhS feedback. Because the central galaxy luminosity functions of individual haloes are strongly peaked around a particular luminosity, the global central galaxy luminosity function is dominated by haloes in a narrow range of mass at each luminosity. (The peak for galaxies in the range $10^{10.5}$ to $10^{11.5}\,h^{-1}\,\mathrm{M}_\odot$ is much broader because these mass scales are affected by PhS, causing particularly strong suppression of galaxy formation at the lower end of the halo mass range. Fig. 1 shows the sharply peaked luminosity distribution of central galaxies in haloes of fixed mass much more clearly.)

The lower left-hand panel of Fig. 4 shows how the satellite and central galaxy luminosity functions combine together to produce the global luminosity function (which is compared to the recent observational determination from the 2dF galaxy redshift survey of Norberg et al. 2002). Brightwards of $M_\mathrm{B} - 5\log h \approx -17$, central galaxies dominate; faintwards of this, the satellite contribution gradually takes over. The faint-end slope of the global luminosity function ends up being intermediate between that of the satellite and central galaxy luminosity functions.

Finally, in the lower right-hand panel of Fig. 4, we show the contribution to the global luminosity function from haloes of different mass. Clearly, the faint-end slope of the global luminosity function is determined in part by the relation between central galaxy luminosity and halo mass, rather than solely by the faint-end slopes of the individual halo luminosity functions. This explains why the global galaxy luminosity function ends up having a flatter faint end than the individual halo luminosity functions. It is interesting to note that both the total and cluster (i.e. $10^{15}\,h^{-1}\,\mathrm{M}_\odot$ halo) luminosity functions are reasonably well-fitted by Schechter functions, albeit with slightly different parameters ($\alpha = -1.31$, $M_* - 5\log h = -20.12$ for the total luminosity function, and $\alpha = -1.45$, $M_* - 5\log h = -19.60$ for the cluster luminosity function; note that the correlation between the parameters α and M_* should be borne in mind when comparing these values).

Diaferio et al. (1999) (see their fig. 1) also used a semi-analytic model to calculate luminosity functions for galaxies in haloes of different masses. While the two models make somewhat different assumptions and adopt different values for key cosmological and galaxy formation parameters, there is good qualitative agreement between the two sets of results (e.g. the trend for increasingly bright galaxies with increasing halo mass). However, there are some notable differences. For example, our model does not produce the strong correlation between the luminosity of the brightest cluster galaxy and cluster mass, which the model of Diaferio et al. found.

4 COMPARISON WITH OBSERVATIONS

Our model predicts marked differences in the luminosity functions of galaxy populations residing in haloes of different mass. Although the detailed form of these predictions depends on the details of our galaxy formation model, the gross differences seen in Fig. 1 are generic predictions that bear directly on basic features of the model, such as hierarchical clustering from CDM initial conditions, gas cooling, photoionization, feedback, etc. Thus, in principle, individual halo luminosity functions provide a strong test of the CDM galaxy formation paradigm. Unfortunately, implementing such a test in practice is difficult, because of the observational complications inherent in identifying galaxies attached to a particular dark matter halo. Gravitational lensing is a powerful and promising tool for detecting dark matter haloes directly, but this technique is still in early stages of development (e.g. Mellier 2002). A less direct (but still useful) approach consists of finding groups and clusters in large redshift catalogues, such as the 2dF and SDSS surveys. When interpreted with the aid of cosmological simulations, it is possible to go some way towards identifying galaxy populations likely to be associated with single haloes (and their subhaloes) of a given mass (Eke et al. 2002). This approach will yield interesting data for comparison with our model predictions in the near future.

In the interim, Trentham & Hodgkin (2002) have provided a limited compilation of observational data. They give estimates of the *B*-band luminosity functions of galaxies in the Local Group and in the Ursa Major, Virgo and Coma clusters. In our model, these systems are expected to reside in haloes of mass similar to those we have simulated, i.e. 10^{12}, 10^{13}, 10^{14} and $10^{15}\,h^{-1}\,\mathrm{M}_\odot$, respectively

(Paper I; Trentham & Hodgkin 2002; Schindler et al. 1999; Geller et al. 1999). We show Trentham & Hodgkin's (2002) compilation in Fig. 1 as open circles. The vertical normalization of the data sets has been adjusted arbitrarily to permit an easier comparison of the shapes and faint-end slopes of the model and observed luminosity functions. For the two most massive systems in the figure, our standard model is in reasonably good agreement with the observed luminosity function, particularly with the faint-end slope.[4] For Ursa Major and the Local Group, the observations indicate a much shallower faint-end slope than is produced by our model. The same discrepancy was noted by Benson et al. (2002b) in their study of the Local Group, and is also apparent in the predicted Local Group luminosity function of Somerville (2002) (which has $\alpha \approx -1.5$ over the range $-20 < M_{\rm V} - 5\log h_{70} < -10$, where h_{70} is the Hubble constant in units of 70 km s^{-1} Mpc^{-1}, for the case of re-ionization at $z = 8$, which is the closest to our own calculations), suggesting that this is a rather robust prediction of galaxy formation models that take PhS into account.

In making the comparison between theory and observations in Fig. 1, one should keep in mind the possibility that the two may not correspond to exactly the same quantity. The theoretical predictions pertain to luminosity functions of the galaxy populations residing within the virial radii of individual dark matter haloes of a given mass. However, there is no guarantee that the samples selected by Trentham & Hodgkin (2002) match this definition. For example, these samples might be contaminated by contributions from secondary haloes, which happen to be close to the dominant halo. This could be particularly important for the Ursa Major cluster, which is a rather diffuse object, and is too large – given its mass – to be associated with a single, virialized halo (according to the standard theoretical definitions). There are ambiguities regarding the Local Group data as well. As argued by Benson et al. (2002b), many dwarfs nominally associated with the Local Group lie beyond the regions which, according to our model, contain the (distinct, but similar-mass) haloes of the Milky Way and M31. In Fig. 1, we show as filled squares the luminosity function of the galaxies that Benson et al. (2002b) identified as lying within the virial radii of either the Milky Way or M31. This luminosity function also shows a fairly flat faint end. Nevertheless, selection effects such as these must be accounted for in a more thorough comparison of theory and observations.

The results of Fig. 1 are re-cast in a different form in Fig. 5, in order to demonstrate the variation in the slope of the luminosity function at different luminosities. Here, we plot the effective luminosity function slope, $\alpha_{\rm eff} = {\rm d}\ln\,\phi/\,{\rm d}\ln L$, as a function of absolute magnitude for all the systems we have simulated. In the model with both PhS and TiL switched off, the luminosity functions in 10^{13} and $10^{14}\,h^{-1}$ M$_\odot$ haloes show an extended region where $\alpha_{\rm eff}$ is almost constant, indicating a power-law luminosity function. With PhS and TiL switched on, the slope is shallower, but the luminosity function is no longer a power law over an appreciable range of magnitudes. For the other mass systems, the luminosity function is nowhere well-described by a power law and, in every case, the inclusion of PhS makes the slope shallower, the effect being larger in the lower-mass systems.

Trentham & Hodgkin (2002) characterized variations in luminosity function shapes using the 'dwarf-to-giant ratio' (i.e. the number of galaxies in some range of faint magnitudes, relative to the number at brighter magnitudes). We compute a similar ratio for our model with PhS and TiL and present the results in Table 1. (We use a slightly different magnitude range for the dwarfs because, for our most massive haloes, our model is not complete over the range of magnitudes considered by Trentham & Hodgkin.)

For the two most massive systems, our model results are in reasonably good accord with the observational determinations, but there are significant discrepancies for the lower-mass systems. This

[4] Our model fails to produce enough of the brightest galaxies, a deficiency which is related to the deficit of mergers in massive systems noted in Section 3.1.

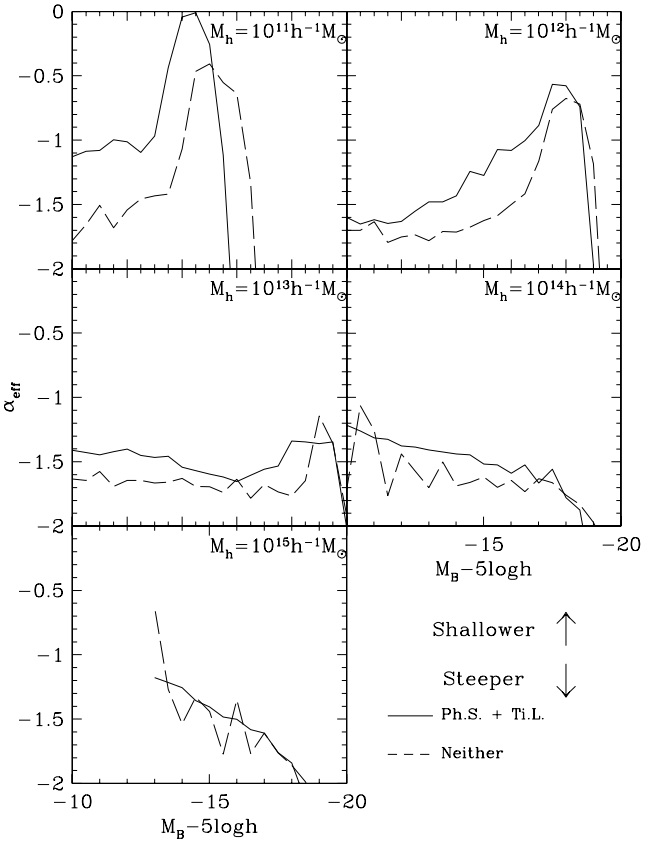

Figure 5. Effective slope of model luminosity functions versus absolute magnitude for galaxy populations residing in haloes of different mass, as indicated in the panels. Solid lines correspond to the model which includes both PhS and TiL, whereas the dashed lines correspond to the model which includes neither.

Table 1. Dwarf-to-giant ratios for galaxy populations in different environments. The magnitude ranges for dwarfs and giants are specified in terms of $M_{\rm B} - 5\log h$ (to maintain consistency with the rest of this work), and correspond to $-16 < M_{\rm B} \leqslant -14$ and $M_{\rm B} \leqslant -16$, respectively, for $h = 0.7$.

Halo mass (h^{-1} M$_\odot$)	$N(-15.23 < M_{\rm B} - 5\log h \leqslant -13.23)/N(M_{\rm B} - 5\log h \leqslant -15.23)$	Observed
10^{11}	0.24	–
10^{12}	0.40	1.25 ± 0.35
10^{13}	1.85	0.74 ± 0.07
10^{14}	1.70	1.47 ± 0.18
10^{15}	1.68	1.40 ± 0.11

analysis verifies and quantifies the comparison between the shapes of the predicted and observed luminosity functions made in Fig. 1.

5 DISCUSSION

Using a semi-analytic model of galaxy formation, we have calculated the luminosity function of galaxy populations contained in haloes of different mass. We find large differences in the shapes of these luminosity functions. In smaller mass haloes ($M \lesssim 10^{13}\,h^{-1}$ $M_\odot$), the luminosity function has a 'hump' at bright magnitudes and a roughly power-law shape at fainter magnitudes. In more massive systems, the shape approaches the familiar Schechter form. These differences reflect the relative contributions of 'central' and 'satellite' galaxies. In the smaller mass haloes, the bright end of the luminosity function is dominated by a single bright galaxy (or, at most, a few of them), and the rest of the population consists predominantly of much fainter satellites. In the more massive haloes, the light is more evenly distributed amongst a large number of galaxies. These differences reflect the complex interplay between the processes that establish the number of haloes and subhaloes of different mass (i.e. hierarchical clustering from CDM initial conditions) and the processes that light them up (i.e. gas cooling, stellar evolution, feedback, etc.).

The global luminosity function is the weighted sum of the luminosity functions of individual haloes. Interestingly, the faint-end slopes of the individual luminosity functions are always steeper than the faint-end slope of the overall luminosity function. This can be understood in terms of the relative contributions to the global function of satellite and central galaxies, the individual halo luminosity functions of which have steep and shallow slopes, respectively. The slope of the global function is intermediate between these two.

Our model of galaxy formation represents an advance over previous models because of its detailed treatment of two processes that are important in establishing the luminosity function: the suppression of galaxy formation in small haloes at high redshift, as a result of the photoionization of the IGM by early generations of galaxies and quasars (PhS); and tidal mass loss from satellite galaxies and subhaloes (TiL).

PhS is particularly relevant at the faint end, and has the net effect of reducing the number of faint galaxies, giving rise to a flatter luminosity function than would otherwise be the case. This outcome is far from obvious. For example, a simplistic model of PhS, in which galaxy formation is assumed not to occur in low-mass haloes below the re-ionization redshift, results, in fact, in a steeper luminosity function (see Appendix A). In reality, the effects of PhS depend on the mass of the halo relative to the time-dependent 'filtering mass' (Gnedin 2000). Galaxies can form even if their haloes have mass smaller than the filtering mass, but with reduced luminosity. The lower the mass of the halo relative to the filtering mass, the fainter the galaxy that forms. The filtering mass is an increasing function of time. Thus, although lower-mass haloes form at higher redshift, the redshift where the filtering mass becomes comparable to their mass is also higher, increasing the fraction of them that experiences strong suppression. These two combined effects *do* tend to produce a flatter luminosity function, but the degree of flattening depends on their detailed balance.[5] As we have shown previously (Paper I; see also Benson et al. 2002b), PhS on its own cannot account for the shallow faint-end slope of the global luminosity function. Other processes, such as feedback from energy released by stellar winds and supernovae, need to be included in order to reduce the number of faint galaxies to the observed level.

Our implementation of PhS is based on the filtering-mass formalism of Gnedin (2000), which predicts that suppression of galaxy formation should become significant in haloes with circular velocities of up to 50 or 60 km s^{-1} at $z \lesssim 5$ (Paper I). This seems reasonably consistent with the earlier results of Thoul & Weinberg (1996) (see also Quinn et al. 1996), who note that '(photoionization suppression) substantially reduces the mass of cooled baryons in systems with circular velocities up to $v_{\rm circ} \approx 50$ km s^{-1}'. In fact, their fig. 7 suggests that the effective filtering mass in their calculation must correspond to circular velocities of from 40 to 50 km s^{-1}. Navarro & Steinmetz (1997), on the other hand, find that a photoionizing background can affect the formation of disc galaxies with circular velocities in the range of from 80 to 200 km s^{-1}, such that the amount of cooled gas in these systems can be reduced by up to 50 per cent. They suggest that this discrepancy with the work of Thoul & Weinberg (1996) may be a result of the idealized nature of the one-dimensional model used in the earlier work, although we note that Weinberg, Hernquist & Katz (1997) found results consistent with Thoul & Weinberg (1996) and inconsistent with Navarro & Steinmetz (1997) in three-dimensional simulations. In conclusion, our implementation of PhS seems to be consistent with the work of Thoul & Weinberg (1996), but its effects are weaker than those found by Navarro & Steinmetz (1997). Clearly, further simulation work is required to establish the strength of this important effect in a conclusive fashion.

The other process that we have treated in detail, TiL, is relevant at the bright end of the luminosity function, particularly in massive haloes harbouring rich clusters. We have shown that our detailed implementation of TiL and satellite orbits prevents the formation of some of the highly luminous and unobserved cluster galaxies that tend to appear in models of galaxy formation that do not take this process into account. Mass loss from satellite haloes, as they spiral into the centre of their host halo, increases the dynamical friction time, thus reducing the merger rate. In Appendix B, we demonstrate that this process is not responsible for preventing the formation of massive central galaxies, but that in fact it is the choice of Coulomb logarithm in the dynamical friction computation which significantly reduces merger rates in clusters.

Keeping in mind the provisos of Section 4 regarding the comparison of theory and data, we find that our model is not a good match to the shape of the luminosity function of the groups and poorer systems that was determined by Trentham & Hodgkin (2002). On the other hand, the model agrees quite well with the data for faint galaxies in the Virgo and Coma clusters. Photoionization is an unavoidable (and effective) mechanism for suppressing galaxy formation. Its effects are greatest in lower-mass systems, and this results qualitatively in the sort of environmental trend identified by Trentham & Hodgkin. However, in low-mass systems, PhS is unable to produce sufficiently flat luminosity functions over the range of magnitudes covered by the data. This failure is related to the discrepancy that we (Benson et al. 2002b, see also Somerville 2002) found previously in our CDM models of the Local Group. Photoionization has a dramatic effect on the overall number of detectable satellites that survive in galactic haloes and can reduce it to the levels seen in the Local Group. However, the satellite luminosity function in the models is considerably steeper than is observed at the faintest magnitudes. This discrepancy contrasts with the success of CDM models

[5] A model in which a fixed cut-off in circular velocity rather than mass is assumed gives results in quite good agreement with our detailed calculations (Paper I).

in reproducing the structural properties of the satellites (Stoehr et al. 2002).

Taking at the face value the discrepancy between our models and the Trentham & Hodgkin (2002) data, we can speculate about the possible root causes of the problem. Leaving aside alternative theories for the nature of the dark matter, or drastic changes to the initial power spectrum of density perturbations that have been specifically designed to reduce the amount of small-scale power (e.g. Colín, Avila-Reese & Valenzuela 2000; Kamionkowski & Liddle 2000), we focus on our modelling of baryonic processes within the context of the CDM cosmogony. This is undoubtedly simplified. For example, the effects of PhS depend upon both the time variation of the filtering mass and the assumed functional form of the suppression, both of which are uncertain. As discussed in the Appendices, reasonable variations in our assumed mass dependence of PhS do not seem capable of producing a sufficiently flat faint-end slope in our models of the Local Group. However, we cannot exclude the possibility that the simplifications made in our model are inappropriate. For example, we neglect the patchy nature of re-ionization, but if low density regions became ionized first, galaxy formation would be suppressed more effectively in these regions, thus producing an environmental dependence of the filtering mass. (Conceivably, there could be environmental variation in the filtering mass, even for homogeneous re-ionization.) Localized PhS triggered, for example, by a nearby galaxy or quasar could also change our model predictions. These processes are poorly understood theoretically, but they could be important in determining the thermodynamic state of the IGM at high redshift.

At a more basic level, there are approximations in our model which might be inappropriate in the regime of low-mass haloes. Thus, although comparisons with N-body/hydrodynamical simulations indicate that our estimates of gas cooling rates are accurate for relatively massive galaxies (Benson et al. 2001a; Yoshida et al. 2002; Helly et al. 2003), it has not been shown that the same is also true for the low-mass galaxies of interest here. Cooling times in these objects can become very short, and this might conceivably render incorrect the assumption of cooling from an initial, spherically symmetric, quasi-hydrostatic equilibrium (Birnboim & Dekel 2003). Further study of gas cooling in low-mass systems is necessary to address this issue. Similarly, it is far from clear that the simple rules for star formation in our model, which are empirical parameterizations tuned to match observational constraints for relatively bright galaxies at low redshifts, remain valid when extrapolated to low-mass galaxies and/or high redshifts. Other physical constraints on star formation that are not currently included in our models may also be important for these low-mass galaxies (e.g. Verde, Oh & Jimenez 2002). Finally, our modelling of feedback is also simplified and, again, its extrapolation to low masses may be unrealistic. Possible sources of feedback which are currently neglected – such as strong outflows or heating by active galactic nuclei – may be important on these scales.

In conclusion, we have found substantial differences in the luminosity functions of galaxies residing in dark matter haloes of different mass. These result from the interplay of a variety of processes which affect the formation of galaxies of different luminosities. Although the exact predictions of our model no doubt depend on its details, the gross differences between the predicted luminosity functions are likely to be generic, and to be present in a broad class of CDM models of galaxy formation. Comparison with existing data gives mixed results: fair agreement with the luminosity functions in Virgo and Coma, but substantial disagreement with the luminosity functions in the Local Group and the Ursa Minor cluster. The much larger and better controlled samples that will be forthcoming from cluster analyses of the 2dF and SDSS galaxy redshift surveys will allow a more comprehensive test of a general class of CDM models of galaxy formation.

ACKNOWLEDGMENTS

We acknowledge the anonymous referee for suggestions which improved this paper. We also thank Simon White for insightful comments and suggestions.

REFERENCES

Balcells M., Peletier R. F., 1994, AJ, 107, 135
Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001a, MNRAS, 320, 261
Benson A. J., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2001b, MNRAS, 327, 1041
Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002a, MNRAS, 333, 156 (Paper I)
Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002b, MNRAS, 333, 177
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, submitted (astro-ph/0302450)
Birnboim Y., Dekel A., 2003, MNRAS, submitted (astro-ph/0302161)
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bower R. G., 1991, MNRAS, 248, 332
Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517
Cole S., 1991, ApJ, 367, 45
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Colín P., Avila-Reese V., Valenzuela O., 2000, ApJ, 542, 622
De Propris R. et al., 2002, MNRAS, 329, 87
Diaferio A., Kauffmann G., Colberg J. M., White S. D. M., 1999, MNRAS, 307 537
Geller M. J., Diaferio A., Kurtz M. J., 1999, ApJ, 517, 23
Gnedin N. Y., 2000, ApJ, 542, 535
Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A. J., Lacey C., Pearce F. R., 2003, MNRAS, 338, 913
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Kamionkowski M., Liddle A. R., 2000, Phys. Rev. Lett., 74, 4525
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Klypin A. A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Kogut A. et al., 2003, ApJ, submitted
Lacey C. G., Cole S., 1993, MNRAS, 262, 627
Lidz A., Hui L., Zaldarriaga M., Scoccimarro R., 2002, ApJ, 579, 491
Mellier Y., 2002, Space Sci. Rev., 100, 73
Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, 19
Navarro J. F., Steinmetz M., 1997, ApJ, 478, 13
Norberg P. et al., 2002, MNRAS, 336, 907
Phillipps S., Shanks T., 1987, MNRAS, 227, 115
Press W. H., Schechter P., 1974, ApJ, 187, 425
Quinn T., Katz N., Efstathiou G., 1996, MNRAS, 278, L49
Schindler S., Binggeli B., Böhringer H., 1999, A&A, 343, 420
Somerville R. S., 2002, ApJ, 572, 23
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Stoehr F., White S. D. M., Tormen G., Springel V., 2002, MNRAS, 335, 84
Thoul A. A., Weinberg D. H., 1996, ApJ, 465, 608
Trentham N., Hodgkin S., 2002, MNRAS, 333, 423
Treu T., Koopmans L. V.E., 2002, MNRAS, 337, 6
Tully R. B., Somerville R. S., Trentham N., Verheijen M. A. W., 2002, ApJ, 569, 573
Verde L., Oh S. P., Jimenez R., 2002, MNRAS, 336, 541
Weinberg D. H., Hernquist L., Katz N., 1997, ApJ, 477, 8

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Yoshida N., Stoehr F., Springel V., White S. D. M., 2002, MNRAS, 335, 762

APPENDIX A: PHOTOIONIZATION AND THE FAINT END OF THE LUMINOSITY FUNCTION

The faint end of the luminosity function directly reflects the action of feedback effects. The two sources of feedback in our model, namely supernova-driven winds and photoionization, are both important in determining the number of faint galaxies. As we discussed in Paper I, photoionization inhibits the formation of faint galaxies for two reasons: it raises the temperature and pressure of the IGM, curtailing the ability of gas to accrete into haloes; and it reduces the rate at which gas in haloes can cool, inhibiting star formation. The first of these processes is the dominant effect. It can be conveniently described in terms of a 'filtering mass', $M_{\rm F}$, defined as the mass of a halo which accretes only half the baryonic mass that it would have accreted in the absence of photoionization. The mass of gas accreted by a halo of mass $M_{\rm h}$ is approximated by the following formula (Gnedin 2000), based on the results of gas-dynamical simulations:

$$M_{\rm gas} = \frac{f_{\rm b} M_{\rm h}}{\left[1 + (2^{1/3} - 1) M_{\rm F}/M_{\rm h}\right]^3}, \tag{A1}$$

where $f_{\rm b}$ is the universal baryon fraction.

There are three different aspects to consider when evaluating the effect of the filtering mass on the faint-end slope of the luminosity function.

(i) Galaxy formation is significantly suppressed in haloes that are less massive than the filtering mass.

(ii) The filtering mass is an increasing function of time (at least over the range of redshifts of interest here).

(iii) For haloes formed at a given redshift, the suppression of galaxy formation is greater the smaller the halo mass is, relative to the filtering mass.

The impact of these three factors will depend on the distribution of formation redshifts of dark matter haloes of different mass in different environments.[6] According to the extended Press–Schechter formalism, it is always true (for a CDM power spectrum) that the typical formation redshift of a halo increases as the mass of the halo decreases. For haloes existing at some earlier redshift which are progenitors of some present-day halo, the typical formation redshift also increases with the mass of the present-day descendent. Further understanding of halo-formation redshift distributions may be gained by considering equation (2.15) of Lacey & Cole (1993), which gives the mass function of progenitor haloes at any redshift, and which we reproduce in a slightly modified form.

$$\frac{{\rm d}N}{{\rm d}M}(z) = \frac{1}{\sqrt{2\pi}} \frac{M_{\rm host}}{M} \frac{\left[\delta_{\rm c}^{(z)} - \delta_{\rm c}^{(0)}\right]}{\left[\sigma^2(M) - \sigma^2(M_{\rm host})\right]^{3/2}} \times \exp\left\{ -\frac{[\delta_{\rm c}^{(z)} - \delta_{\rm c}^{(0)}]^2}{2\left[\sigma^2(M) - \sigma^2(M_{\rm host})\right]} \right\} \frac{{\rm d}\sigma^2}{{\rm d}M}, \tag{A2}$$

where $\delta_{\rm c}^{(z)}$ is the critical value of the linear-theory fractional overdensity for collapse at redshift z, M is the mass of the progenitor halo at that redshift, $\sigma^2(M)$ is the variance of the linear fractional overdensity in a sphere of mass M, and $M_{\rm host}$ is the mass of the halo into which the progenitor has been incorporated by the present day.

Consider galaxies living in relatively low-mass haloes at the present day (in which reside most of the galaxies that make up the faint end of the luminosity function). The relative numbers of haloes of masses M_1 and $M_2 (< M_1)$ at some particular redshift depends upon $\sigma^2(M_1) - \sigma^2(M_{\rm host})$ and $\sigma^2(M_2) - \sigma^2(M_{\rm host})$. As $M_{\rm host}$ increases, $\sigma^2(M_{\rm host})$ decreases for a CDM power spectrum. Thus, for sufficiently large values of $M_{\rm host}$, the relative numbers of haloes of masses M_1 and M_2 tend to a fixed value. For values of $M_{\rm host}$ comparable to those of M_1, the relative numbers of these haloes become a very strong function of $M_{\rm host}$, and the number of mass M_1 haloes is exponentially suppressed, as indicated by equation (A2).

These theoretical expectations are clearly manifest in Fig. A1, where we plot the positions of galaxies in the halo-mass versus formation-redshift plane. All of the galaxies we consider at $z = 0$ are satellites in a larger host halo (of mass $M_{\rm host}$) because of the ranges of luminosity and $M_{\rm host}$ we have chosen to consider. Therefore, we plot the formation redshift and mass for the halo in which the satellite formed. We adopt the definition of formation redshift of Cole et al. (2000) – namely that a halo is assumed to be newly formed if its mass exceeds twice the mass of each of its progenitors at their own formation time – and plot the halo mass at the formation redshift (it may have increased afterwards, owing to continued accretion and merging). The position of a halo in this plane, relative to the redshift-dependent filtering mass (shown as a solid curve), determines the degree of PhS it experiences. We examine faint galaxies in two luminosity ranges: $-13 < M_{\rm B} - 5 \log h < -12$ (dots) and $-15 < M_{\rm B} - 5 \log h < -14$ (small squares).

Consider first a model with no PhS (left-hand panels). In this case, faint galaxies currently hosted in a $10^{14}\,h^{-1}\,{\rm M}_\odot$ cluster halo (lower panel) have a much more extended range of formation redshifts than galaxies in a $10^{12}\,h^{-1}\,{\rm M}_\odot$ Milky Way-type halo (upper panel). The fainter galaxies located in lower-mass haloes have a more extended range of formation redshifts than their brighter counterparts, although this difference is rather small. The values of α (calculated from the two luminosity ranges considered) for the two populations, given in each panel, indicate steep slopes in both environments. The filtering mass has no influence on these results, of course, but is shown to indicate which of these haloes might still be expected to form a galaxy when PhS is switched on.

In the right-hand panels, we plot results for a model with PhS. The effect of the filtering mass can now be seen clearly: haloes below the solid line form galaxies much less efficiently than before. The values of α given in the panels indicate that some flattening of the luminosity function has occurred, but this effect is small.

Let us now examine the importance of these various processes, using simplified models of PhS. Fig. A2 shows luminosity functions in different mass haloes constructed using the simplified models described below, together with our standard model of PhS (heavy solid line) and a model with no PhS (thin solid line).

Model A. In the simplest scenario, suppose that the filtering mass jumps abruptly from zero to $M_{\rm F}$ at redshift $z_{\rm reion}$, and that any halo of smaller mass forming at lower redshift makes no galaxy at all, whereas galaxy formation in other haloes proceeds unchanged. In this case, the luminosity function should be unaffected at bright magnitudes (because these galaxies form in haloes which are more massive than $M_{\rm F}$), should plummet sharply at a magnitude corresponding to a halo mass equal to $M_{\rm F}$, and should rise again to fainter magnitudes (as these galaxies inhabit lower-mass haloes which

[6] By 'environment', we mean the mass of the halo into which the earlier halo has been incorporated by the present day.

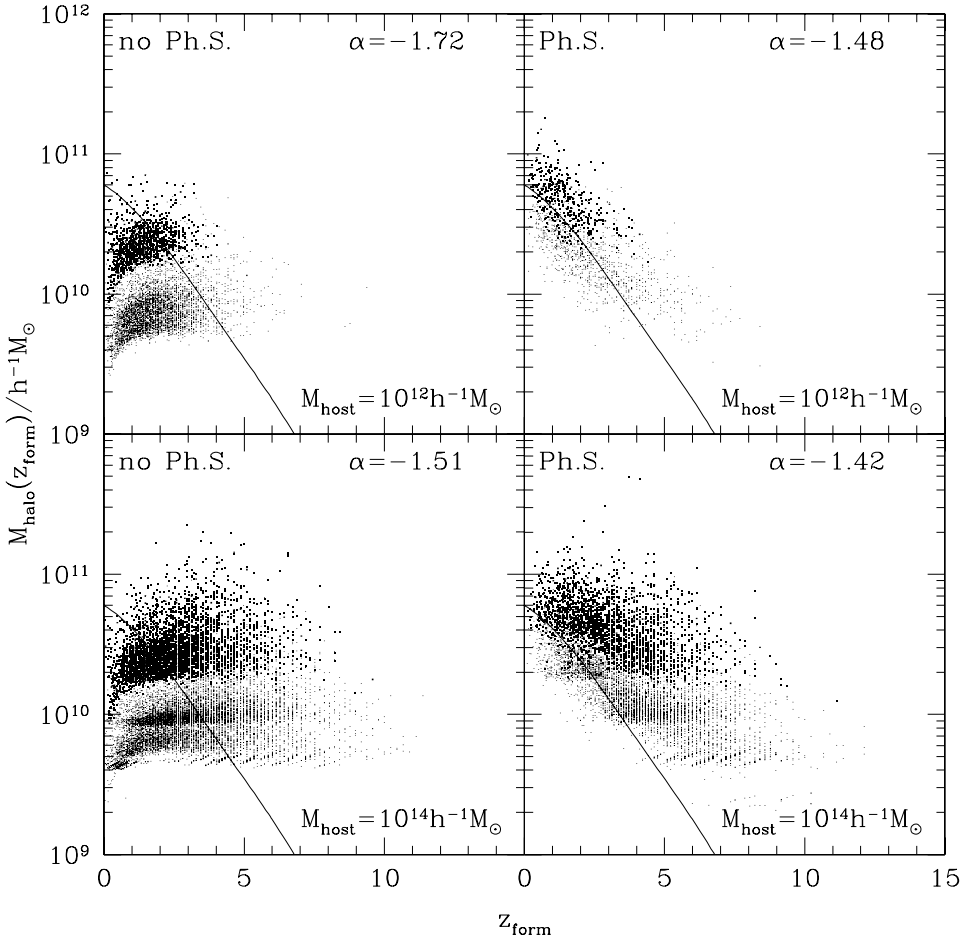

Figure A1. Positions of galaxies in the halo-mass/formation-redshift plane. Upper and lower panels show results for galaxies which reside in $10^{12}\,h^{-1}$ $M_\odot$ and $10^{14}\,h^{-1}$ $M_\odot$ haloes at $z = 0$, respectively. Left- and right-hand panels show results for models ignoring PhS and including PhS, respectively. The solid line shows the filtering mass as a function of redshift. Dots indicate galaxies in the range $-13 < M_B - 5\log h < -12$, whereas small squares indicate galaxies in the range $-15 < M_B - 5\log h < -14$. An estimate of α based on these two magnitude bins is shown in each panel.

typically form at redshifts greater than z_{reion}). On the basis of equation (A2), the slope of this rise should be large for low-M_{host} haloes, because these have the greatest variation in progenitor number with mass at z_{reion}, and should approach some fixed value for large values of M_{host}. This behaviour is borne out when we apply this simple prescription to our semi-analytic models, using values of $M_F = 3 \times 10^{10}\,h^{-1}$ $M_\odot$ (comparable to the value of M_F at the redshifts where the majority of the faint galaxies are formed in our standard model) and $z_{reion} = 6.5$ (the redshift of re-ionization in our standard model). The result is shown by the dot–long-dashed lines in Fig. A2 – suppression is greatest for small values of M_{host}, tends to steepen the faint end of the luminosity function, and causes the most steepening for the lowest values of M_{host}. This model produces a discontinuity in the luminosity function and is a poor approximation to the full treatment.

Model B. At the next level of sophistication, we take into account the variation of the filtering mass with redshift as predicted by our model, but retain the simple prescription wherein any halo forming with mass below the filtering mass forms no galaxy at all. Because M_F increases with time, galaxy formation in lower-mass haloes is suppressed earlier – an effect which should tend to flatten luminosity function slopes. Because haloes of a given mass form earlier in larger values of M_{host}, we still expect the effects of filtering to be weaker for clusters than for the Local Group. At a given magnitude, this process would simply reduce the number of galaxies. However, the detailed consequences for the faint-end slope are now harder to anticipate, because they depend on exactly how M_F varies with redshift. Applying this prescription to our model (long-dashed lines in Fig. A2), we still find a discontinuity in the luminosity function (although this is generally much weaker than in model A, because M_F varies smoothly with time), which is flattened relative to that of the model without any PhS. Furthermore, the slopes are significantly flatter than in case A, indicating that the time variation of M_F largely counteracts the steepening of the luminosity function produced by the variation in halo-formation distributions. No environmental variation is introduced by these processes. The resulting luminosity functions are quite similar to those from our full PhS calculations, although the level of suppression is somewhat greater. Again, the effects are largest for the lowest values of M_{host}.

Model C. Alternatively, we can consider a prescription which has no time variation in M_F, but in which, instead, the degree of suppression varies smoothly with halo mass, as prescribed by equation (A1).[7] We find that this model (dot–short-dashed lines)

[7] Somerville (2002) considered a similar model, but kept the circular velocity corresponding to M_F constant, resulting in a filtering mass that increased with time. This prescription, also adopted by Bullock et al. (2000), provides

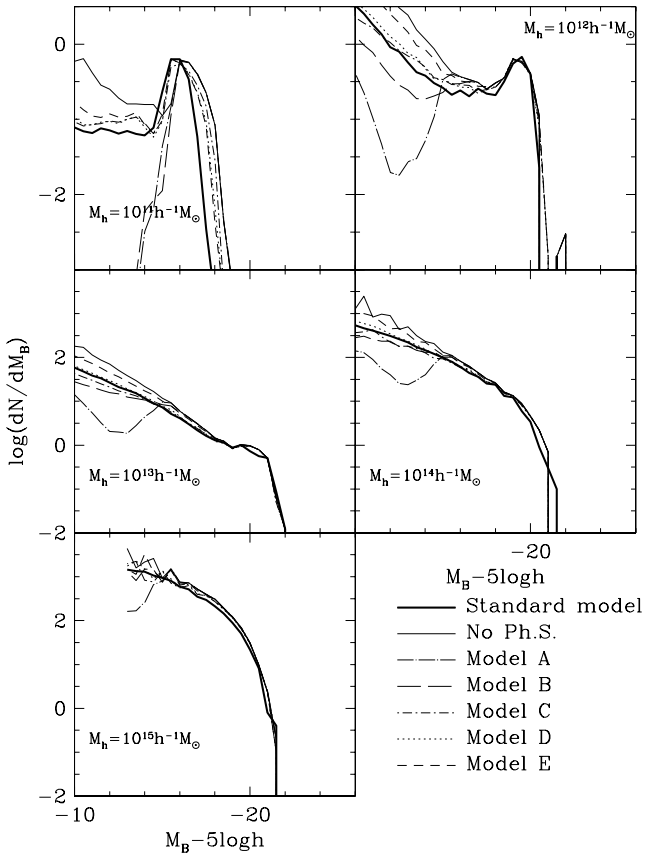

Figure A2. *B*-band galaxy luminosity functions in haloes of different mass. Each panel shows the mean predicted model luminosity function in an ensemble of dark matter haloes of mass given in each panel (ranging from haloes containing small galaxies to rich clusters). All models include the effects of tidal limitation. The heavy solid lines show results from our complete model of PhS, whereas thin solid lines show a model with no PhS. Dot–long-dashed lines have a filtering mass which jumps abruptly from zero to $3 \times 10^{10}\,h^{-1}\,\mathrm{M}_\odot$ at $z = 6.5$, and have a sharp suppression at that mass at lower redshifts (Model A). Long-dashed lines have a time-varying filtering mass, but have a sharp suppression at that mass (Model B). The dot–short-dashed line shows a model with a fixed filtering mass, but with a smooth variation of suppression with halo mass (Model C). Dotted lines are a simplified model with time-varying filtering mass, and with a smooth variation of suppression with halo mass (Model D). Short-dashed lines are the same, but with an extra-smooth variation of suppression with halo mass (Model E).

produces luminosity functions quite close to those from our full PhS calculations, for suitable choices of M_F and z_reion.

Model D. We next consider a model in which M_F varies with time and in which the degree of suppression varies smoothly with halo mass. This then differs from our full PhS calculations only in neglecting suppression, as a result of photoheating by the ionizing background. In this model, photoionization redistributes galaxies – which previously had a given magnitude – over a range of fainter magnitudes. At a given magnitude, some galaxies are lost as a result of suppression; however, new ones also appear as brighter galaxies are partially suppressed, thus becoming fainter. The net effect is therefore difficult to judge, because it depends on the time variation of the filtering mass and on the previous shape of the luminosity function. However, we can safely say that the effect is still expected to be larger for lower values of M_host. Applying this prescription, we reproduce the results of our full PhS calculation rather well (dotted lines in Fig. A2).

Model E. A smoother variation of suppression with halo mass would result in less flattening of the faint-end slope. For example, the short-dashed lines in Fig. A2 show results for a suppression of the form

$$M_\mathrm{gas} = \frac{f_\mathrm{b} M_\mathrm{h}}{1 + M_\mathrm{F}/M_\mathrm{h}}, \tag{A3}$$

with the same variation of filtering mass with redshift as in our standard model.

In conclusion, a simple model in which galaxy formation at $z < z_\mathrm{reion}$ is entirely suppressed below some fixed mass scale grossly overestimates the effects of PhS, and produces very different luminosity functions from those predicted by our full calculation. Including *either* a smoothly varying degree of suppression with mass *or* a time-varying filtering mass produces better agreement with our full calculations, although the smooth variation with mass seems more important. The degree of suppression (as characterized by the reduction in amplitude of the LF) for a given time dependence in M_F depends on the functional form used to characterize the variation of suppression with halo mass. The sharper the transition from weak suppression to strong, the greater the net effect on the luminosity function. Smoother transitions tend to produce steeper faint-end slopes, but even the sharpest transition possible only produces a flat luminosity function in the very lowest mass systems we consider, and significant environmental variation occurs only when mass changes discontinuously with time.

APPENDIX B: DIFFERENCES BETWEEN GALAXY MERGER MODELS

It is interesting to compare in some detail the results of the two rather different merging models used in this work. We employ the detailed model of galaxy merging of Benson et al. (2002b) as standard, which tracks the orbit of each individual satellite galaxy and its dark matter halo (assumed to have the NFW density profile), accounting for mass loss due to tidal forces. We contrast this with the simpler model of Cole et al. (2000), in which merging times are calculated, assuming an isothermal halo model and neglecting tidal mass loss.

In the model of Cole et al. (2000), the time taken for a galaxy to merge can be written as

$$\tau_\mathrm{Cole} = \frac{f_\mathrm{orb}}{2\pi B(1) \ln \Lambda} \frac{M_\mathrm{halo}}{M_\mathrm{sat}} \tau_\mathrm{dyn} = f_\mathrm{orb} t_\mathrm{scl}, \tag{B1}$$

where $B(1) \approx 0.427$, M_sat is the total mass of the satellite galaxy (including its dark matter halo), M_halo is the mass of the halo in which it orbits, τ_dyn is the dynamical time of the halo, $\ln \Lambda$ is a Coulomb logarithm which Cole et al. (2000) took to be $\Lambda = M_\mathrm{halo}/M_\mathrm{sat}$, f_orb is a factor containing the dependence on the orbital parameters of the satellite and t_scl is defined by the above relation. Fig. B1 shows the distribution of merging times in this model, expressed in units of t_scl (dotted histogram). (Specifically, this is the distribution for mergers involving subhaloes of mass greater than $10^{10}\,h^{-1}\,\mathrm{M}_\odot$ occurring during the assembly of a $10^{15}\,h^{-1}\,\mathrm{M}_\odot$ cluster. Similar distributions are found for haloes of other masses.) As may be seen from the figure, Cole et al. (2000) assumed a lognormal distribution for f_orb, motivated by results from *N*-body simulations.

Also shown in Fig. B1 is the distribution of merging times (also in units of t_scl) in the detailed model used in this work. This peaks in

a better match to the results of our standard model. Here, however, we are interested only in exploring simplified models.

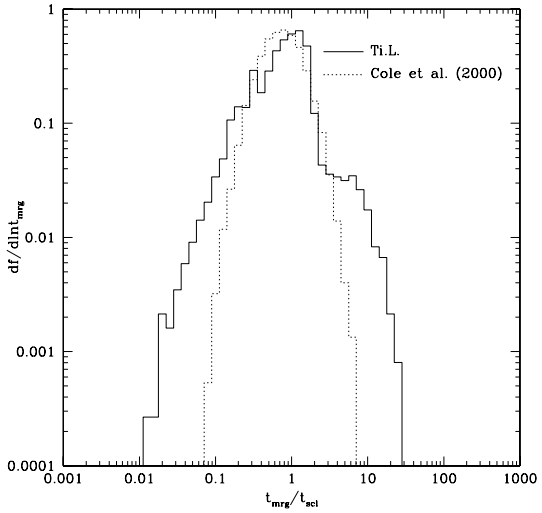

Figure B1. The distribution of merging times [expressed in units of $t_{\rm scl}$, as defined in equation (B1)] for mergers of haloes more massive than $10^{10}\,h^{-1}$ $\mathrm{M}_{\odot}$, occurring during the formation of a $10^{15}\,h^{-1}$ $\mathrm{M}_{\odot}$ cluster. The solid histogram shows results from the model employed in this work, whereas the dotted histogram shows results from the model of Cole et al. (2000).

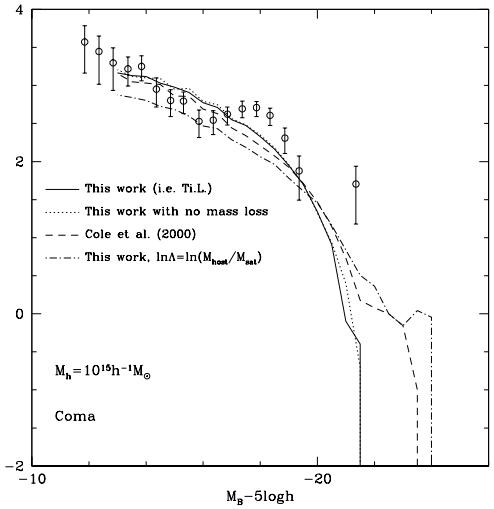

Figure B2. *B*-band luminosity functions of galaxies in $10^{15}\,h^{-1}$ $\mathrm{M}_{\odot}$ clusters. The plot gives the mean number of galaxies per magnitude, per halo. Open circles show the observed luminosity function from the compilation of Trentham & Hodgkin (2002), normalized arbitrarily, to permit easier comparison of the shape with the models. The solid line shows results using our detailed model of merging, including TiL, whereas the dashed line shows results from the model of Cole et al. (2000). The dotted line shows results from the detailed model of merging, but with the effects of mass loss switched off. Finally, the dot-dashed line corresponds to the detailed model, but with the Coulomb logarithm set equal to ln $M_{\rm host}/M_{\rm sat}$.

the same place as the distribution of the model of Cole et al. The two have similar mean and median merging times, but the distribution from the detailed model is significantly broader. In this case, those objects with extremely long merging time-scales have penetrated the central regions of their host halo (where the central galaxy often dominates the potential) and have experienced significant tidal mass loss. This greatly weakens the dynamical friction they experience, thereby extending their merging time.

In Fig. 1, we showed that the Cole et al. (2000) model forms very massive galaxies at cluster centres, but the detailed model suppresses their formation. A possible cause of this difference is the effect of mass loss in the detailed model, which, as we have just seen, extends the merging time of some galaxies. This explanation can be tested simply by switching off mass loss in the detailed model, while retaining all other aspects of the calculation. Fig. B2 shows the resulting luminosity function – it is essentially indistinguishable from the detailed model including mass loss.

In fact, it is the Coulomb logarithm, ln Λ, that is responsible for the differences between the two models. Cole et al. (2000) chose $\Lambda = M_{\rm halo}/M_{\rm sat}$; in our detailed model, we use $\Lambda = r\sigma^2/\mathrm{G}M_{\rm sat}$ instead, where r is the current orbital radius of the satellite and σ is the velocity dispersion of the halo at that point. This choice gives better fits to the orbital evolution of satellites in N-body simulations. Typically, Λ is a factor of $\approx$4 smaller than the simple estimate of Cole et al., corresponding to a difference of $\approx$1.4 in the Coulomb logarithms. For most mergers, where ln Λ is significantly larger than *1*, this makes little difference to the merging times. In clusters, however, merging is only likely for rather massive satellites, for which $t_{\rm scl}$ is sufficiently small. A large value of $M_{\rm sat}$ corresponds to a small Coulomb logarithm, giving rise to large fractional differences between the two calculations. Considerably longer merger times are then obtained in the detailed model than in the model of Cole et al.

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