Galaxy formation in the Planck cosmology - I. Matching the observed evolution of star formation rates, colours and stellar masses

Article (Published Version)

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **451,** 2663–2680 (2015) doi:10.1093/mnras/stv705

Galaxy formation in the *Planck* cosmology – I. Matching the observed evolution of star formation rates, colours and stellar masses

Bruno M. B. Henriques,[1]★ Simon D. M. White,[1] Peter A. Thomas,[2] Raul Angulo,[3] Qi Guo,[4] Gerard Lemson,[1] Volker Springel[5,6] and Roderik Overzier[7]

[1]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85741 Garching b. München, Germany*
[2]*Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK*
[3]*Centro de Estudios de Física del Cosmos de Aragón, Plaza San Juan 1, Planta-2, E-44001 Teruel, Spain*
[4]*Partner Group of the Max-Planck-Institut für Astrophysik, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China*
[5]*Heidelberger Institut für Theoretische Studien, Schloss-Wolfsbrunnenweg 35, D-69118 Heidelberg, Germany*
[6]*Zentrum für Astronomie der Universität Heidelberg, ARI, Mönchhofstr. 12-14, D-69120 Heidelberg, Germany*
[7]*Observatório Nacional/MCTI, Rua José Cristino, 77. CEP 20921-400, São Cristóvão, Rio de Janeiro-RJ, Brazil*

Accepted 2015 March 27. Received 2015 March 27; in original form 2014 October 3

ABSTRACT

We have updated the Munich galaxy formation model to the *Planck* first-year cosmology, while modifying the treatment of baryonic processes to reproduce recent data on the abundance and passive fractions of galaxies from $z = 3$ down to $z = 0$. Matching these more extensive and more precise observational results requires us to delay the reincorporation of wind ejecta, to lower the surface density threshold for turning cold gas into stars, to eliminate ram-pressure stripping in haloes less massive than $\sim 10^{14}\,\mathrm{M}_{\odot}$, and to modify our model for radio mode feedback. These changes cure the most obvious failings of our previous models, namely the overly early formation of low-mass galaxies and the overly large fraction of them that are passive at late times. The new model is calibrated to reproduce the observed evolution both of the stellar mass function and of the distribution of star formation rate at each stellar mass. Massive galaxies ($\log M_{\star}/\mathrm{M}_{\odot} \geq 11.0$) assemble most of their mass before $z = 1$ and are predominantly old and passive at $z = 0$, while lower mass galaxies assemble later and, for $\log M_{\star}/\mathrm{M}_{\odot} \leq 9.5$, are still predominantly blue and star forming at $z = 0$. This phenomenological but physically based model allows the observations to be interpreted in terms of the efficiency of the various processes that control the formation and evolution of galaxies as a function of their stellar mass, gas content, environment and time.

Key words: methods: analytical – methods: statistical – galaxies: evolution – galaxies: formation – galaxies: high-redshift.

1 INTRODUCTION

Galaxy formation theory has developed dramatically over the last three decades. Λ cold dark matter (ΛCDM) has been established as the standard model for cosmological structure formation, and its parameters have been increasingly tightly constrained by observations. In parallel, simulations of galaxy formation within this standard model have grown in complexity in order to treat more accurately the many baryonic processes that impact the evolution of the galaxy population. Semi-analytic modelling is a particular simulation method which is optimized to connect the observed properties of the galaxy population – abundances, scaling relations, clustering and their evolution with redshift – to the astrophysical processes that drive the formation and evolution of individual galaxies (e.g. White 1989; Cole 1991; Lacey & Silk 1991; White & Frenk 1991; Kauffmann, White & Guiderdoni 1993; Cole et al. 1994; Kauffmann et al. 1999; Somerville & Primack 1999; Springel et al. 2001, 2005; Hatton et al. 2003; Kang et al. 2005; Lu et al. 2011; Benson 2012). Simple phenomenological descriptions of the relevant processes are needed, each typically involving uncertain efficiency and scaling parameters. These must be determined by comparison with observation or with more detailed simulations. As the range and quality of observational data have increased, so has the number of processes that must be included to model them adequately, and hence the number of adjustable parameters. In recent years, robust statistical methods have been introduced in order to sample the resulting high-dimensional parameter spaces and to determine the regions that are consistent with specific observational data sets. This development began with the work of Kampakoglou, Trotta &

★ E-mail: bhenriques@mpa-garching.mpg.de

Silk (2008) and Henriques et al. (2009) and has since been extended to a wide range of models and sampling methods (Benson & Bower 2010; Bower et al. 2010; Henriques & Thomas 2010; Lu et al. 2011, 2012; Henriques et al. 2013; Mutch, Poole & Croton 2013; Benson 2014; Ruiz et al. 2015).

Semi-analytic modelling is designed to interpret the statistical properties of the galaxy population, and particular emphasis has always been placed on local galaxies for which abundances, scaling relations and clustering are best determined. The sharp high-mass cut-off in the observed stellar mass function can be explained by invoking efficient feedback from central black holes (Benson et al. 2003; Granato et al. 2004; Bower et al. 2006; Croton et al. 2006) but the properties of low-mass galaxies, where star formation is limited by strong stellar feedback, remain poorly reproduced by even the most recent models. Typically, the fraction of low-mass galaxies ($8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$) that are no longer star forming is substantially overestimated at low redshift, as is their high-redshift abundance (Weinmann et al. 2006b, 2012; Henriques, Bertone & Thomas 2008; Fontanot et al. 2009; Guo et al. 2011, 2013; Henriques et al. 2012; Hirschmann et al. 2014). These difficulties have become more prominent in recent years as observational surveys at high redshift have improved. Massive galaxies ($\log M_\star/\mathrm{M}_\odot \geq 11.0$) seem to have assembled most of their mass by $z = 1$, while the abundance of low-mass galaxies grows substantially at later times (Fontana et al. 2006; Faber et al. 2007; Pozzetti et al. 2010; Marchesini et al. 2009, 2010; Ilbert et al. 2010, 2013; Muzzin et al. 2013). Furthermore, vigorous star formation is almost ubiquitous in nearby low-mass galaxies, while most massive galaxies are currently red and appear to have formed few stars since $z \sim 1$ (e.g. Kauffmann et al. 2003; Baldry et al. 2004; Brinchmann et al. 2004; Thomas et al. 2005; Arnouts et al. 2007; Drory et al. 2009; Peng et al. 2010).

In a recent paper (Henriques et al. 2013), we showed that delaying the reincorporation of gas ejected by supernova-driven winds can shift the mass assembly of dwarf galaxies to lower redshift without significantly affecting massive systems. Here, we include this change and modify additional aspects of the Munich model which affect star formation in low-mass galaxies. With respect to Guo et al. (2011), we decrease the cold gas density threshold for star formation and we eliminate ram-pressure stripping in all but the most massive dark haloes, a more drastic version of the modification advocated for similar reasons by Font et al. (2008). These changes increase the fraction of low-mass satellite galaxies that are blue at low redshift. The fraction of low-mass centrals that are blue is also increased substantially by the delayed reincorporation of wind ejecta, so together these changes produce a dwarf population which is predominantly blue at $z = 0$. We also change our model for AGN feedback to make it more efficient at late times. This is needed to ensure that galaxies around the knee of the stellar mass function are predominantly quenched by $z = 0$ despite growing significantly in number at $z \leq 2$. Finally, we adjust the underlying cosmological parameters of our model to correspond to the first-year *Planck* results.

By fully sampling the predicted properties for a high-dimensional space of model parameters, we demonstrate that the observed evolution of the abundance of galaxies can be reproduced as a function of stellar mass *and* specific star formation rate (sSFR) by plausible representations of the relevant astrophysics within the standard ΛCDM structure formation model. To achieve an adequate representation of current observations, it is necessary to change the modelling of a number of astrophysical processes – as we will show, our earlier models cannot fit these new data for *any* values of their parameters. Furthermore, these new data provide moderately tight constraints on all the parameters of the updated model leaving no major degeneracies. The fact that the updated model provides a good fit to the observed stellar mass functions and passive fractions over the full redshift range, $0 \leq z \leq 3$, is a significant success. It also predicts many other properties of galaxies (e.g. sizes, morphologies, gas fractions, clustering) which were not explicitly used in setting its parameters and which can be used to test it and to refine it further. Throughout this paper, we will refer to all galaxy properties computed by the model, that are not directly used as inputs, as model predictions. These include the properties used to define the best-fitting set of parameters.

In a companion paper (Henriques et al., in preparation, hereafter Paper IV), we analyse the processes that cause galaxies to migrate between the active and passive populations, comparing with observationally derived quenching efficiencies as a function of stellar mass, halo mass, local density and clustering, and focusing in particular on the data presented by Peng et al. (2010), Wetzel, Tinker & Conroy (2012) and Kauffmann et al. (2013). In future work, we will further test our model by comparing additional properties to available observations.

Detailed predictions for this new model are made publicly available with this paper.[1] These include snapshot and light-cone outputs with pre-computed magnitudes for various stellar population models and a wide range of broad-band filters, as well as simulated optical spectra and full star formation histories (see Section 3.3 below). In the following section, we briefly describe modifications in the current model with respect to previous versions. The extensive Supplementary Material presents a full description of the treatment of astrophysical processes. Section 3 describes how we set parameters in our new model. Section 4 compares its predictions to our calibrating observational data sets, while Section 5 compares to additional data. We conclude with a short summary of our main results.

2 UPDATES TO GALAXY FORMATION MODELLING

In this section, we describe how the treatment of astrophysical processes within the Munich galaxy formation model has changed since the last publicly released catalogues which were based on Guo et al. (2011). A complete description of the current model, including more details on the newly implemented procedures, can be found in the Supplementary Material which is published along with this paper on the arXiv and is available online.[2] The changes were driven by known problems with earlier versions of this (and many other) model. Specifically, the overly early ($z \geq 2$) build-up of low-mass galaxies ($8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$) is eliminated by changing the time-scales for gas to be reincorporated after ejection in a wind; the related problems that low-mass galaxies end up too red, too old and too clustered at $z = 0$ are eliminated by additionally lowering the cold gas density threshold for star formation and removing ram-pressure effects on satellites in low-mass groups ($\log M_{200}/\mathrm{M}_\odot < 14.0$); the problem that too many massive galaxies ($\log M_\star/\mathrm{M}_\odot \sim 11.0$) continue to form stars at low redshift is eliminated by assuming radio mode feedback to scale with global, rather

[1] The Munich models of galaxy formation can be found at http://www.mpa-garching.mpg.de/millennium

[2] A complete description of the model presented in this paper, as well as all the theoretical functional forms plotted, are available for download at http://galformod.mpa-garching.mpg.de/public/henriques2014a/

than local, halo properties. In addition, we rescale the Millennium Simulations to represent the cosmology preferred by the first-year results from *Planck*.

2.1 Simulations and cosmology

The Munich model of galaxy formation has been implemented on two simulations of the evolution of dark matter structure. Combined, the Millennium (Springel et al. 2005) and Millennium-II (Boylan-Kolchin et al. 2009) simulations provide a dynamic range of five orders of magnitude in stellar mass ($10^{7.0}\,\mathrm{M}_\odot < M_\star < 10^{12}\,\mathrm{M}_\odot$), resolving the smallest galaxies observed at $z = 0$ while also sampling the largest clusters. Over the stellar mass range $10^{9.5}\,\mathrm{M}_\odot < M_\star < 10^{11}\,\mathrm{M}_\odot$, Guo et al. (2011) found good numerical convergence between the two simulations, both for the abundance of galaxies and for their mass- and colour-dependent clustering. A similar level of convergence is found for most of the properties discussed in this paper, and throughout we will use the Millennium Simulation to derive properties for galaxies more massive than $10^{9.5}\,\mathrm{M}_\odot$ and the Millennium-II to derive properties at lower mass.

We use the technique of Angulo & White (2010), as updated by Angulo & Hilbert (2015), to scale the evolution of dark matter structure predicted by the Millennium and Millennium-II Simulations to the recently published *Planck* cosmology. Specifically, the cosmological parameters we adopt from Planck Collaboration XVI (2014) are: $\sigma_8 = 0.829$, $H_0 = 67.3\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$, $\Omega_\Lambda = 0.685$, $\Omega_\mathrm{m} = 0.315$, $\Omega_\mathrm{b} = 0.0487$ ($f_\mathrm{b} = 0.155$) and $n = 0.96$. As can be seen in Fig. 1, structural growth for this new set of cosmological parameters is as close to growth in the original Millennium (*WMAP1*) cosmology as to that in most of the cosmologies implied by subsequent *WMAP* releases. For example, in terms of structure formation, the *WMAP5* cosmology chosen for the Bolshoi (Klypin, Trujillo-Gomez & Primack 2011) and Multi-Dark (Prada et al. 2012) simulations differs from the currently preferred *Planck* cosmology by twice as much as the original *WMAP1* cosmology of the Millennium Simulations in terms of the scaling required (∼9 per cent in box size and ∼29 per cent in particle mass compared to ∼4 per cent in box size and ∼17 per cent in particle mass), but only by half as much in terms of the time offset ($\Delta z = 0.07$ rather than $\Delta z = 0.12$). However, as shown by Wang et al. (2008), Guo et al. (2013) and Fontanot et al. (2012), the differences in cosmological parameters between all these modern determinations have a much smaller effect on galaxy properties than the uncertainties in galaxy formation physics. As a result, the change in cosmology has little impact in our conclusions.

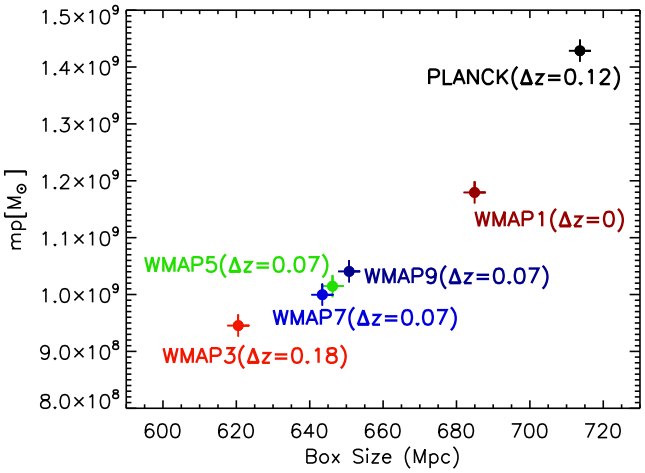

Figure 1. The box size and particle mass required for the Millennium Simulation to give a good representation of structure formation in different underlying cosmologies. Numbers in brackets give the difference in redshift between $z = 0$ in the target cosmology and in the *WMAP1* cosmology of the original simulation. Hence, to represent the *Planck* cosmology for this paper, the box size and the particle mass have to be scaled up by about 4 and 21 per cent, respectively. The $z = 0.12$ output of the original simulation becomes the new $z = 0$.

2.2 Astrophysical modelling

2.2.1 Reincorporating wind ejecta

A number of recent studies have found that both semi-analytic and hydrodynamic simulations of the evolution of the galaxy population tend to form low-mass galaxies, $8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$, too early ($z \geq 2$), with the result that they are substantially more abundant than observed at early times (Fontanot et al. 2009; Weinmann et al. 2012; Lu et al. 2014). In Henriques et al. (2013), we showed that a possible solution to this problem is to couple strong feedback with a significantly increased delay in the time for gas to be reincorporated after ejection in a galactic wind. This substantially reduces the growth of low-mass systems at high redshift, which is then compensated by enhanced growth between $z = 2$ and 0 as the ejected gas finally returns and fuels star formation. The result is a much better match to the observed evolution of the stellar mass function.

In the current work, we adopt the specific implementation proposed by Henriques et al. (2013), where the time-scale for material to be reincorporated scales directly with halo mass and is independent of redshift. A very similar dependence on parameters was found by Oppenheimer & Davé (2008) and Oppenheimer et al. (2010) in their cosmological hydrodynamics simulations, which gave reasonable fits to the evolution of the low-mass stellar mass function. In practice, this assumption means that wind ejecta from low-mass haloes take a long time to return, while ejecta from massive systems are almost immediately reincorporated. Since the return rate that we assume depends on the *current* mass of a halo rather than on its mass at the time of ejection, substantial halo mass growth can considerably accelerate the reincorporation of previously ejected material.

Firmani, Avila-Reese & Rodríguez-Puebla (2010) implemented a similar dependence of gas re-accretion efficiencies in an attempt to obtain down-sizing of star formation rates in analytic models of star-forming disc galaxies in a ΛCDM universe. They argued that re-accretion alone cannot explain observed trends, which appeared to require inclusion of additional processes such as AGN feedback and a detailed model for cooling. Recently, Mitchell et al. (2014) and White, Somerville & Ferguson (2015) implemented the Henriques et al. (2013) reincorporation scheme in their own semi-analytic models. While the former show results consistent with ours, the latter found that it did not prevent the early build-up of low-mass galaxies. This presumably indicates the importance of fully exploring how this process interacts with details of the implementation of all the other physical processes that shape galaxy populations. Neistein & Weinmann (2010) and Wang, Weinmann & Neistein (2012) obtained a similar delay in the build-up of galaxies with $8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$ at early times as in our new reincorporation model by choosing appropriate scalings of cooling efficiency with halo mass and redshift. We do not allow similar freedom in our own model, since we consider the cooling of diffuse gas to be one of the better understood aspects of galaxy formation, and the scalings

adopted here agree reasonably well with hydrodynamic simulations (Forcada-Miro & White 1997; Benson et al. 2001; Yoshida et al. 2002).

Recently completed hydrosimulations (Vogelsberger et al. 2014; Schaye et al. 2015), have significantly improved the level of agreement between the evolution of the stellar mass function predicted by these methods and observations (Furlong et al. 2014; Genel et al. 2014; Crain et al. 2015). It would therefore be interesting to compare the gas reincorporation time-scales in these simulations with our assumptions and we plan to do so in future work.

2.2.2 *Star formation thresholds*

A related problem which was obvious in previous versions of the Munich model (and in many other galaxy formation models) is the fact that a large fraction of simulated galaxies with $8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$ are no longer forming stars by $z = 0$ and hence are red, whereas Sloan Digital Sky Survey (SDSS) data indicate that the great majority of real low-mass galaxies are, in fact, blue (Guo et al. 2011, 2013). The delayed build-up of these low-mass galaxies produced by our new reincorporation model reduces the discrepancy between model and observations but does not remove it completely. While low-mass central galaxies re-absorb gas at late times, thereby fuelling continued star formation, environmental effects remove the gas supply of low-mass satellites which then turn red relatively quickly.

A study of the properties of these satellites reveals that while much of their gas is removed by interactions with their host halo, they still retain a significant amount of cold gas in their discs. This gas is not transformed into stars, however, because the model adopts a gas surface density threshold for star formation. Recent observations suggest that any such threshold is lower than previously thought, and that star formation is better modelled as being linked to molecular rather than to total gas surface density, hence depending on the processes that convert HI to H_2 (Bigiel et al. 2008; Leroy et al. 2008). Detailed semi-analytic models for atomic to molecular gas conversion have been developed by Fu et al. (2012, 2013) and Lagos et al. (2011) and will be included in future versions of the Munich model. For this paper, we just note that, in practice, stars can form from gas with surface density below our previous standard threshold, so we simply decrease this threshold by approximately a factor of 2. Satellites can then consume a larger fraction of their cold gas and so keep forming stars for longer periods.

2.2.3 *Environmental effects*

Although lowering the threshold for star formation decreases the number of red satellites, these remain considerably more numerous than observed. This seems to indicate that the environmental suppression of star formation is too strong in our model. A variety of processes affect satellite galaxies, but a detailed analysis indicates that the removal of hot gas reservoirs by ram-pressure stripping is already sufficient to quench star formation in satellites to below the observed levels. Indeed, suppressing this effect entirely (with no change to other processes) results in a good match to the observed fraction of passive satellites (Kang & van den Bosch 2008; Weinmann et al. 2010). Since ram-pressure stripping is *observed* to occur in rich clusters where there are substantial X-ray emitting hot gas atmospheres, we elect to retain the process in haloes with $M_{200c} > M_{\mathrm{R.P.}} = 10^{14}\,\mathrm{M}_\odot$. In lower mass groups and clusters, X-ray data show significantly lower hot gas fractions, at least in the inner regions (see Sun 2012 for a review on the subject). Given that characteristic orbital velocities are also substantially lower in such systems, it is plausible that stripping effects should be less important there, and indeed Font et al. (2008) already advocated such a reduction with respect to the findings of direct simulations of the stripping process, noting that this significantly improved the colours of dwarf galaxies in their own galaxy formation model.

The analytic equations derived by McCarthy et al. (2008) suggest that in order for satellite galaxies to retain enough hot gas to fuel continued star formation at the levels observed in low-mass groups, the hot gas surface density of the group must be 30 times smaller than that of the satellite. This requires feedback to produce large reductions in gas content and concentration in groups. More detailed work is needed to see if this can be consistent with observations and hence explain the observed lack of environmental influence on star formation. Here, we decide simply to remove the problem assuming that ram-pressure stripping does not extend to low-mass groups, including the corresponding mass threshold as an additional free parameter in our Markov Chain Monte Carlo (MCMC) analysis.

In the preferred model of this paper, we thus leave all other environmental effects unchanged. In consequence, a significant fraction of satellites are quenched, even in groups where ram-pressure stripping is inactive. As in Guo et al. (2011, 2013), infall of new material stops immediately when a halo falls within R_{200c} of a larger system and tidal stripping of hot gas is then turned on. The latter parallels the stripping of dark matter from the subhalo, a process which the original *N*-simulation followed explicitly. The hot gas reservoir is removed completely once a satellite loses its own halo and becomes an 'orphan'. The model assumes such orphans are unable to retain gas ejected by supernova (SN) feedback which is moved to the hot halo of the galaxy group. Tidal forces can completely disrupt the stellar and cold gas components of orphan galaxies, which are then added to the intra-cluster light and the hot gas atmosphere of the group/cluster, respectively. Our treatment of all these processes is fully described in the Supplementary Material.

In Paper IV, we will analyse the effects of our environmental physics assumptions by comparing our model in detail both with traditional autocorrelation function measurements of clustering, as in Guo et al. (2011), and with more recent studies of the dependence of galaxy properties on the mass of the halo they occupy and on their position within it, in particular, the studies by Peng et al. (2010, 2012), Wetzel et al. (2012) and Kauffmann et al. (2013). We conclude that our new model matches most but not all of the observed trends highlighted in these papers.

2.2.4 *Radio mode feedback*

The new model of Section 2.2.1 for the reincorporation of wind ejecta ensures that the abundance of galaxies below the knee of the stellar mass function can grow substantially from $z = 2$ to 0, as observed, reflecting considerable growth in mass of the individual systems at late times. With the radio mode feedback model of Guo et al. (2011), which was taken directly from Croton et al. (2006), this late-time growth results in too many star-forming galaxies of stellar mass close to $M_\star$ and above at $z = 0$ (see Henriques et al. 2013, and Fig. 5 below). This undesirable change can be eliminated by modifying the AGN feedback model to suppress cooling and star formation more effectively at low redshift than did the original model.

We follow the methodology of Henriques et al. (2013) and include the exponents which determine the scaling of radio mode

feedback with halo mass and redshift as additional parameters in an MCMC comparison with multiredshift observations of stellar mass functions and passive fractions. We find that a relatively small change to the radio mode feedback model of Croton et al. (2006) can indeed restore the cross-over between predominantly star forming and predominantly passive galaxies to the observed stellar mass while retaining the substantial growth in abundance of low-mass galaxies at low redshift. Specifically, taking the heating rate to be $\dot{E} \propto M_{\rm BH} M_{\rm hot}$ rather than $\dot{E} \propto M_{\rm BH} M_{\rm hot} H(z)$ (the form used in Croton et al. 2006 and Guo et al. 2011) substantially improves the match to observation. The abundance of galaxies at $M_\star$ and below grows significantly at $z < 2$, yet star formation is quenched in most $M_\star$ objects by $z = 0$.

The analysis presented later in this paper and especially in Paper IV, shows that this simple model for radio mode feedback quenches star formation in intermediate and high-mass galaxies at rates and in locations that are roughly consistent with observations. Nevertheless, even though our new implementation ensures that most massive galaxies are red and are dominated by old populations at late times, there is an indication that some may still be more actively star forming than observed.

In addition to its scaling with hot gas properties, we have changed how AGN radio mode feedback from satellite galaxies affects the hot gas of the host halo. For satellites inside R_{200}, any AGN energy left over after offsetting the cooling of the hot gas of the satellite is used to offset cooling of the hot gas of the host halo. This change was introduced to fix a numerical artefact identified in the dark matter merger trees. In one group of galaxies, a small object ($\log M_\star/{\rm M}_\odot \sim 9.5$) centred on a low-mass subhalo ($\log M_{200}/{\rm M}_\odot \sim 12.0$) becomes the central galaxy of a massive friends-of-friends (FOF) group. Lacking a massive central black hole, and suddenly increasing its assigned hot gas by orders of magnitude (from $\log M/{\rm M}_\odot \sim 11.0$ to $\log M/{\rm M}_\odot \sim 14.0$), this galaxy then experiences catastrophic growth through cooling to become the most massive galaxy in the whole simulation. This problem was already present in previous models but is accentuated in the current version, where the delayed build-up of galaxies means that their black holes grow preferentially at late time, increasing the likelihood that a low-mass satellite galaxy does not yet have a massive black hole.

Our modification ensures that, for the group where this happened, the energy released by black holes in other cluster galaxies suppresses excessive cooling of the intracluster medium. We have checked that this change only impacts the desired object.

2.2.5 *Stellar populations and dust model*

Galaxy formation models most naturally predict the evolution of physical properties of galaxies such as stellar masses, gas fractions, star formation rates, sizes, characteristic velocities and morphologies. To convert these into observables such as photometric luminosities, sizes and colours requires additional modelling of galactic stellar populations and dust content in order to calculate the emission and absorption of light at specific wavelengths, based on the age and metallicity distribution of the stars and their spatial distribution relative to interstellar dust.

Henriques et al. (2011, 2012) showed that, at least in the context of the Munich model, stellar populations that have significant emission from TP-AGB stars in intermediate age populations (e.g. Maraston 2005; Charlot & Bruzual 2007) appear required to reconcile the observed evolution of the stellar mass and K-band luminosity functions. For this paper, we therefore use Maraston (2005) with a Chabrier (2003) initial mass function (IMF) as our default stellar population model. We have tested that the as yet unpublished Charlot & Bruzual (2007) model gives similar results for all properties analysed in this paper. As part of our data release, we will also include light-cones with luminosities computed using the older Bruzual & Charlot (2003) model, though we caution that the rest-frame near-IR luminosities appear to be significantly underestimated in some situations in this case.

In producing photometric catalogues, we also adjusted the high-redshift conversion of gas/metals into dust. In Guo et al. (2011) this fraction of gas/metals in dust was assumed to scale as $(1 + z)^{-0.4}$ whereas we now adopt $(1 + z)^{-1}$. This change ensures that galaxies at very high redshift have very little extinction, as observed, and results in luminosity functions for Lyman-break galaxies at $z \geq 5$ that match *Hubble Space Telescope* (*HST*) data quite well. These effects will be discussed in Paper III of this series (Clay et al., 2015). All quantities presented in this paper are unaffected by this change. We emphasize that this is a purely phenomenological fix and may be required simply to rectify the impact of our excessively high metallicities at early times.

3 SETTING PARAMETERS

3.1 Fiducial data and MCMC procedures

The modelling framework set out above scales a high-resolution dark matter only simulation of cosmic structure formation to a specific desired cosmology (here the *Planck* cosmology) and implements a set of simplified phenomenological models that can be applied in post-processing to describe the processes that affect the formation and evolution of the galaxies themselves. Specifying a particular model requires choosing values for the parameters that determine the efficiency and the scaling of these processes. Some can be set to sufficient accuracy from physical arguments or by comparison with more detailed simulations, but the rest must be determined by fitting to some fiducial set of observational data.

There has been substantial recent progress in determining the abundance of galaxies as a function of stellar mass and star formation activity. Multiple surveys now appear to give broadly consistent results out to $z \geq 3$. For this paper, we have therefore elected to use the observed stellar mass function of galaxies at $z = 0$, 1, 2 and 3, together with the observed fraction of passive ('red') galaxies as a function of stellar mass at $z = 0$, 0.4, 1, 2 and 3 to determine the parameters of our 'best' model. As in previous work (Henriques et al. 2009, 2013; Henriques & Thomas 2010), we combine results from all modern determinations of each quantity into a single 'representative' data set with error bars representing the scatter between determinations as well as the statistical uncertainties in the individual measurements. We show these fiducial observations in Figs 2 and 5 together with the best-fitting model found by our MCMC exploration of parameter space. These plots are discussed in more detail below. The original observations underlying our fiducial data sets are shown in Appendix A.

Our MCMC procedures also parallel those of our earlier work (e.g. Henriques et al. 2013). In particular, we calculate a figure of merit, in effect an approximate likelihood, for each model given our representation of the observations by assuming that each blue point is independently sampled from a Gaussian with mean given by the model and variance given by the sum of an 'observational' variance indicated by the error bar and a 'theoretical' variance which we take to be negligible for the stellar mass functions and to be 0.025^2

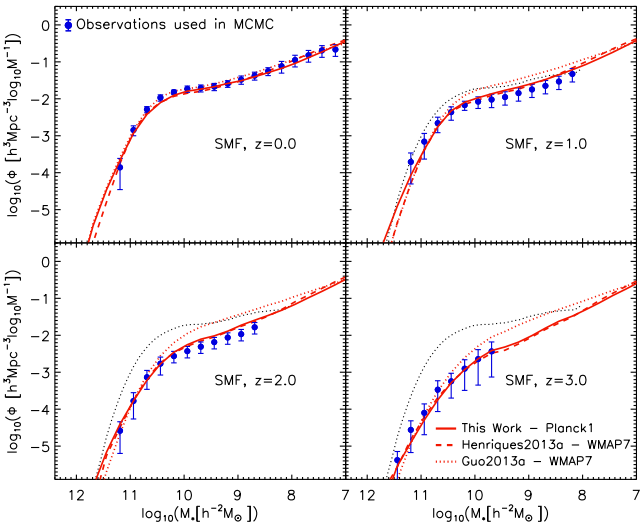

Figure 2. Evolution of the stellar mass function from $z = 3$ to 0. Lines show predictions from our new model (solid red), from Henriques et al. (2013, dashed red) and from Guo et al. (2013, dotted red). These should be compared with the blue symbols with error bars which represent the combined observational data which we use to constrain the MCMC. As described in Appendix A, the data sets we combine include SDSS (Baldry, Glazebrook & Driver 2008; Li & White 2009) and GAMA (Baldry et al. 2012) at $z \sim 0$, and Marchesini et al. (2009), Spitzer-COSMOS (Ilbert et al. 2010), NEWFIRM (Marchesini et al. 2010), COSMOS (Domínguez Sánchez et al. 2011), ULTRAVISTA (Ilbert et al. 2013; Muzzin et al. 2013) and ZFOURGE (Tomczak et al. 2014) at higher redshift. The $z = 0$ combined stellar mass function is repeated at higher redshift as a dotted black line. Here and in all subsequent plots, predicted stellar masses have been convolved with a Gaussian in $\log M_\star$, with width $0.08 \times (1 + z)$, in order to account for the uncertainties in observational stellar mass determinations.

for the red fractions. Note that our analysis neglects any sampling covariance between the different data points. Together with the fact that our error bars are subjective assessments of uncertainty, rather than direct estimates of sampling variance, this means that our MCMC procedure does not give statistically rigorous estimates of parameters and their uncertainties. Instead, it should be interpreted as an efficient method for delineating the region of parameter space which produces acceptable representations of the observational data.

3.2 Best-fitting model

For this paper, we sample a significantly larger parameter space than in Henriques et al. (2013), including almost all the explicit free parameters in our semi-analytic model. An exception is the threshold between major and minor mergers. We find that there is some tension between the value of this parameter required to match observations of the fraction of red galaxies (one of our primary constraints) and that required to match galaxy morphologies. We slightly compromise the agreement with the observed red fraction of galaxies at $z = 1$ by fixing $R_{\rm merge}$ at 0.1 in order to obtain reasonable morphologies for massive galaxies at $z = 0$. The best-fitting values for the 17 parameters included in our MCMC sampling are tabulated together with their marginalized 2σ ranges in the Supplementary Material which also contains, for reference, all the equations that define our galaxy formation model, as well as a more detailed analysis of the allowed confidence regions for its parameters and a comparison of their values with previous models. The predictions of our new best-fitting model are shown as solid red lines in Figs 2 and 5 where it can be seen to represent the fiducial observations quite well. Although our updated cosmology, together with the changes in our modelling of ejecta reincorporation, of star formation, of ram-pressure stripping and of AGN feedback induce shifts by factors of several in a number of parameters as compared, for example, to Guo et al. (2011, 2013), all parameters still remain well within their physically plausible ranges.

3.3 Public catalogues

With the completion of this paper, we will provide public access to catalogues of data from our new 'best-fitting' model. These will have the same format as previous data releases and will be available through the same SQL interface (Lemson & Virgo Consortium 2006). As before, they will include data both for individual snapshots of the Millennium and Millennium-II simulations, and for light-cones constructed from the larger volume simulation. As part of the GALFORMOD project, we have recently expanded the functionality of these archives: a new interface allows asynchronous queries to the data base which avoid timeouts. In terms of the available predictions, the public catalogues will follow previous releases and will include the possibility to link galaxy and dark matter halo data bases, hence to analyse jointly a wide range of galaxy and dark halo properties. As previously for the Henriques et al. (2012) light-cones, we will include galaxy magnitudes for a large number of broad-band filters and for several stellar population models. A further extension of previous releases is the inclusion of full star formation histories for every galaxy (fully described in Paper II of this series, Shamshiri et al. 2015), allowing the construction of synthetic spectra, using an arbitrary dust model, in post-processing.

4 FIT TO THE FIDUCIAL OBSERVATIONS

In this section, we compare our best-fitting new model in more detail with the observational data which are used as constraints in our MCMC: the evolution from $z = 3$ to 0 of the stellar mass function, and the fraction of galaxies that are 'red' (i.e. no longer actively star forming) as a function of stellar mass and redshift. We pay particular attention to the fact that low-mass objects ($8.0 \leq \log M_\star/{\rm M}_\odot \leq 9.5$) assemble most of their mass and form most of their stars later than massive systems. This trend is often referred to as 'downsizing' and seems superficially to contradict the hierarchical growth of structure expected in a ΛCDM cosmology. It has been known for some time, however, that this reversal reflects the baryonic physics of galaxy formation (e.g. De Lucia et al. 2006), and the particular implementation in this paper reproduces the observed trends in considerable detail. Dwarf galaxies form most of their stars late when material ejected in early winds is finally reincorporated. Massive galaxies ($\log M_\star/{\rm M}_\odot \geq 11.0$) are barely affected by SN feedback and grow quickly at high redshift, turning off when their central black holes grow big enough to suppress further accretion and so quench star formation.

In this paper, we focus on the evolution of global galaxy properties, mainly stellar masses and star formation rates, distinguishing between trends at low and high stellar mass, and for passive and actively star-forming galaxies. In Paper IV, we will look in more detail at trends as a function of environment and their evolution with redshift. Both here and in Paper IV it will turn out that the changes in the treatment of reincorporation, of star formation and of

ram-pressure stripping are particularly important for dwarf galaxies, while our new treatment of AGN feedback has substantial effects only for massive objects.

In all the following sections, model stellar masses have been convolved with a Gaussian in $\log M_\star$, with width increasing with redshift, in order to mimic the uncertainties in observational stellar mass determinations. The scatter is assumed to have similar redshift dependence to that found by Ilbert et al. (2013), but larger size [a Gaussian with dispersion $0.08 \times (1+z)$ instead of $0.04 \times (1+z)$]. We note that this is slightly larger than assumed by Behroozi, Wechsler & Conroy (2013a) ($0.07 + 0.04z$) and still considerably smaller than found by Conroy, Gunn & White (2009) even neglecting IMF uncertainties (0.3 at $z = 0$ and 0.6 at $z = 2$). Similar uncertainties were found by Pforr, Maraston & Tonini (2012) and Mitchell et al. (2013) when testing spectral energy distribution (SED) fitting methods using galaxies from semi-analytic models. As discussed in Section 2.1 we use the Millennium Simulation to derive properties for galaxies more massive than $10^{9.5}\,\mathrm{M}_\odot$ and the Millennium-II to derive properties at lower mass.

4.1 Evolution of the stellar mass function – 'down-sizing' in action

Fig. 2 shows galaxy stellar mass functions from $z = 3$ to 0. Results from our new model (solid red lines) are compared with results from Henriques et al. (2013, dashed red lines) and from Guo et al. (2013, dotted red lines), both of which assume a *WMAP7* cosmology. As explained in Appendix A, the blue symbols with error bars are the representation of the observations used in our MCMC procedures and were produced by combining a number of recent observational studies in an attempt to estimate systematic uncertainties in the constraints. Data from the various individual observational studies are plotted independently in Fig. A1. With respect to Henriques et al. (2013), we include new data sets from the UltraVISTA survey (Ilbert et al. 2013; Muzzin et al. 2013), the ZFOURGE survey (Tomczak et al. 2014) and the GAMA survey for the lowest redshift bin (Baldry et al. 2012). The new surveys are considerably deeper than previous data and have relatively large areas, allowing us to extend our fiducial data set to lower stellar masses with relatively low statistical uncertainty. When needed, we follow Domínguez Sánchez et al. (2011) and apply a correction of $\Delta M_\star = -0.14$ to go from Bruzual & Charlot (2003) to Maraston (2005) stellar populations at $z \geq 1$.

After adjusting the parameters of our model to fit these new data, the predicted abundance of galaxies as a function of stellar mass and redshift is almost identical in our *Planck* cosmology and in the *WMAP7* cosmology of Henriques et al. (2013). As noted by Wang et al. (2008) and Guo et al. (2013), the uncertainties in galaxy formation physics produce much larger differences in galaxy masses than any change of cosmology within the currently allowed range. As already shown in Henriques et al. (2013), the longer timescales for gas reincorporation in low-mass haloes combine with stronger SN feedback to reduce the abundance of galaxies with $8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$ at high redshift. The return of ejected gas at later times results in a significant build-up at $z \leq 1$, as required by the observational data. This late return does not drive similar low-redshift growth in massive galaxies ($\log M_\star/\mathrm{M}_\odot \geq 11.0$) because AGN feedback and less efficient cooling result in the production of hot gas atmospheres rather than further star formation in these systems. Overall, our new galaxy formation model (like the model of Henriques et al. 2013 before it) is able to explain the observed evolution in the stellar mass function over the last 80 per cent of cosmic history and over the full mass range constrained by observations.

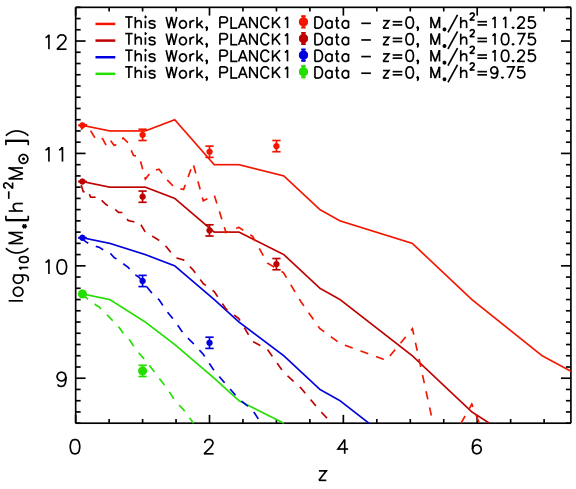

Figure 3. Mass growth with redshift of galaxies with different $z = 0$ mass $M_{\star,0}$. The symbols with error bars are derived from the observed stellar mass functions shown in Fig. 2 by matching the cumulative abundance of galaxies $n(M_\star, z)$ at each redshift z to $n(M_{\star,0}, 0)$. Solid lines show model predictions derived using this same abundance-matching scheme, while dashed lines show the mean stellar mass at redshift z of the most massive progenitors of all galaxies that have stellar mass $M_{\star,0}$ at $z = 0$.

The trend for lower mass galaxies to form their stars at later times than high-mass ones (i.e. 'down-sizing') can be seen more directly in Fig. 3. Here, model results for the growth in mass of individual galaxies are shown for systems that have $\log_{10}(M_\star[h^{-2}\,\mathrm{M}_\odot])$ 9.75, 10.25, 10.75 and 11.25 at $z = 0$. Dashed curves give the mean stellar mass of the most massive progenitor as a function of redshift. From least massive to most massive final galaxy, the growth factors since $z = 2.5$ are 28, 18, 9 and 7, respectively, showing that giant galaxies indeed have a larger fraction of their mass in place at high redshift than dwarfs.

It is not, of course, possible to observe directly the growth histories of individual galaxies, but a number of recent papers suggest that these can be inferred by assuming that the typical progenitor mass of galaxies that have stellar mass $M_{\star,0}$ today, $M_\star(z, M_{\star,0})$, can be inferred from the 'abundance-matching' equation $n(M_\star, z) = n(M_{\star,0}, 0)$, where $n(M_\star, z)$ is the comoving abundance at z of galaxies with stellar mass exceeding $M_\star$ (e.g. van Dokkum et al. 2010; Brammer et al. 2011; Papovich et al. 2011). This argument neglects the scatter in assembly history among galaxies of given $M_{\star,0}$ as well as the non-conservation of galaxies due to mergers. The solid curves in Fig. 3 show the result of deriving growth histories in this way in our model. It is clear that abundance matching leads to significant underestimation of the true amount of growth, giving factors of 9, 6, 4 and 3 since $z = 2.5$, two to three times smaller than the correct values. These results, for the mass growth of progenitors, are consistent with the findings of Behroozi et al. (2013b) and represent a larger difference between the true mass growth and the cumulative number densities method than found for the evolution of descendants (Leja, van Dokkum & Franx 2013). Points with error bars in Fig. 3 show the result of applying this abundance-matching argument to our representation of the observational data (i.e. the blue points with error bars in Fig. 2) and hence should be compared with the solid curves. There is good agreement in the two higher mass bins, but some disagreement at lower masses.

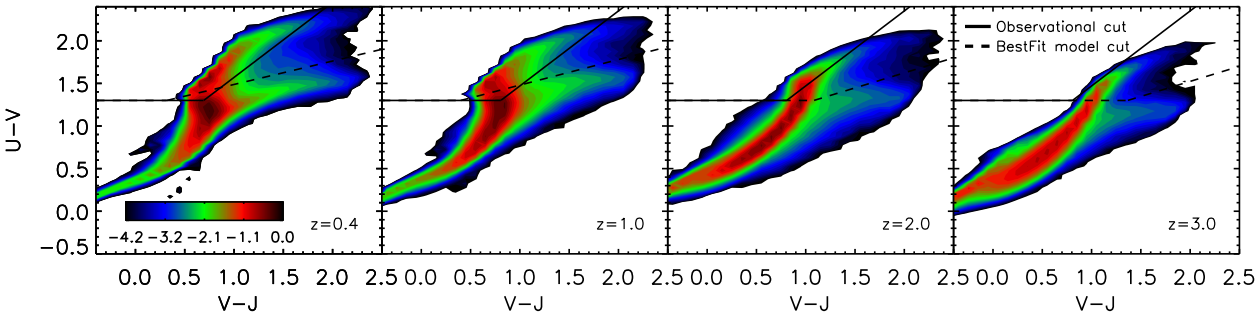

Figure 4. The $U - V$ versus $V - J$ colours of galaxies in our model. The solid line represents the observational separation into red and blue subpopulations, while the dashed line separates active and passive galaxies in our best-fitting model (see text). The colour-scale contours represent the normalized number density of galaxies in logarithmically spaced bins (from dark red high-density to dark blue low-density contours).

4.2 The red galaxy fraction – passive systems dominate at high mass and low redshift

The second set of constraints we use when setting the parameters of our new model are the fractions of passive galaxies as a function of stellar mass and redshift. These are obtained from observed stellar mass functions split by colour into active and passive systems, and are defined as the ratio of the estimated abundance of 'red' to 'red+blue' systems for stellar mass bins where estimates are available for both types of object. We start by detailing the methods used to separate red and blue galaxies in the model. Later in this section, we will compare the best-fitting model with our adopted constraints (with error bars determined by propagating the errors we assigned the original abundances) which are shown as blue points in Fig. 5. The original colour-separated mass functions from which they were derived are shown in Fig. 7. We prefer to use these passive fractions in our MCMC sampling, rather than the colour-separated mass functions themselves, in order to separate more cleanly the constraints coming from abundances from those coming from star formation activity.

A number of criteria have been proposed to separate star forming from passive galaxies. Most involve a cut in colour or in inferred star formation rate. Recent observational studies have advocated separating galaxies using both optical and optical to near-infrared colours since this allows truly passive systems to be distinguished from dusty star-forming galaxies. Here, we use rest-frame $U - V$ and $V - J$ colours, as proposed by Muzzin et al. (2013) and also used by Ilbert et al. (2013) and Tomczak et al. (2014). Fig. 4 shows the $U - V$ versus $V - J$ rest-frame colour distribution for model galaxies at four different redshifts. A clear separation can be seen in all panels between passive galaxies (top middle), blue star-forming galaxies (bottom left) and dusty star-forming galaxies (middle right). In the observational samples, truly passive galaxies are found to lie above and to the left of the solid lines in Fig. 4. The separation is not, however, in exactly the same place in our model, presumably because of short-comings in our stellar population synthesis and dust modelling. We therefore modified the separatrix at red $V - J$ colour from the solid to the dashed line when estimating passive fractions in our model. The appropriate split appears unambiguous, except, perhaps, at $z = 3$. This division in colour correctly separates star forming and passive galaxies in the model whereas using the original observational separation would result in passive simulated galaxies at $z \geq 2$ being identified as 'dusty star-formers' simply because our populations synthesis/dust modelling gets their colours wrong.

At $z = 0$, we apply a cut in the magnitude–colour plane in order to match the criteria used for the observations with which we compare (Baldry et al. 2004). We again adopt a slightly different cut than in the observations when separating active and passive galaxies in the model: $u - r = 1.85 - 0.075 \times \tanh((\mathrm{M_r} + 18.07)/1.09)$. All the cuts are fixed at the separation between the two populations in the best-fitting model and unchanged during the MCMC process. In practice, the cut is obtained iteratively by initially running a shorter MCMC with a colour cut based on the separation from an earlier best fit. For the model, we have tested that using a colour–magnitude cut, a colour–colour cut or a cut in sSFR yields very similar red and blue fractions across all redshifts. Our conclusions are thus insensitive to this choice.

Our modifications to the physics of AGN radio mode feedback, to ram-pressure stripping, and to the threshold for star formation have a substantial impact on red galaxy fractions. This can be seen in Fig. 5 which compares our representation of the observational data with predictions from our new model (the solid red lines), from the model of Henriques et al. (2013, the dashed red lines) and from that of Guo et al. (2013, the dotted red lines). Given that these observational data were used as constraints in our MCMC sampling, it is no surprise that the new model provides the best fit of the three. In particular, it predicts a significantly lower fraction of passive low-mass galaxies ($8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$) at $z \leq 1$ than both earlier models, and a substantially higher fraction of passive giant galaxies ($\log M_\star/\mathrm{M}_\odot \geq 10.5$) at low redshift than the Henriques et al. (2013) model (though still somewhat fewer than observed and than in the Guo et al. (2013) model at $z = 0.4$ and 1). Passive objects are only 20 per cent of the observed $z = 0$ population for $\log M_\star/\mathrm{M}_\odot \leq 9.5$, and our model is now consistent with this small number.

In order to understand how our updated modelling has altered passive galaxy fractions, it is instructive to split galaxies of each type into central and satellite systems. Fig. 6 shows the stellar mass functions of active and passive galaxies at $z = 0$ as blue and red solid curves, respectively. Passive galaxies dominate the population for stellar masses above $3 \times 10^{10}\mathrm{M}_\odot$, while actively star-forming galaxies are dominant at lower masses. Dashed and dotted curves then separate each of these functions into the contributions from central and satellite galaxies, respectively.[3] Central galaxies switch from predominantly star forming to predominantly passive at $4 \times 10^{10}\mathrm{M}_\odot$ and this transition, induced by our AGN feedback prescription, is quite sharp. At $10^{11}\mathrm{M}_\odot$ more than 80 per cent of centrals are passive, whereas at $10^{10}\mathrm{M}_\odot$ this is the case for about 20 per cent; below $2 \times 10^{9}\mathrm{M}_\odot$ only about 2 per cent of centrals are

[3] Note that our definition of satellite here includes any galaxy that is not centred on the main subhalo of its FOF dark matter halo.

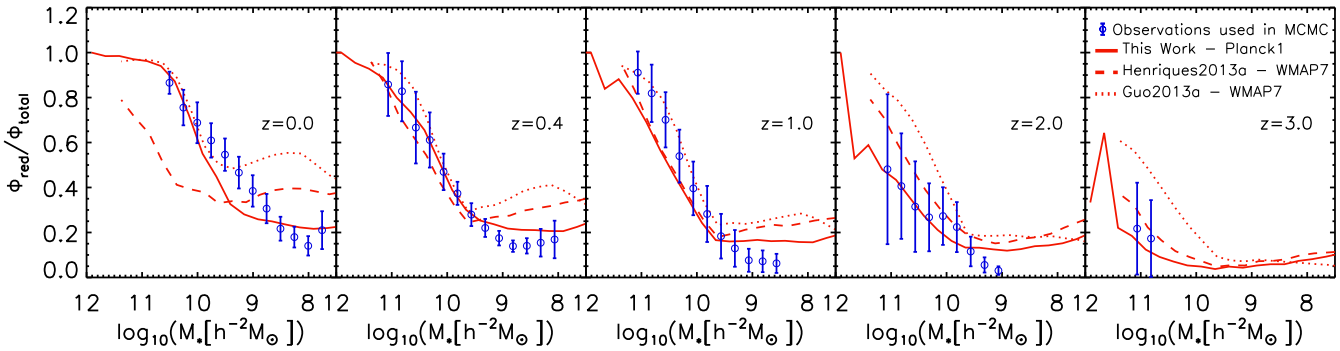

Figure 5. The evolution of the fraction of red (passive) galaxies as a function of stellar mass and redshift. Predictions from our new model (solid red lines), from that of Henriques et al. (2013, dashed red lines) and from that of Guo et al. (2013, dotted red lines) are compared with observed red fractions (blue points with error bars). These were obtained by dividing the stellar mass function of red galaxies by the sum of the red and blue stellar mass functions shown in Fig. 7. Error bars come from straightforward propagation of the errors shown in Fig. 7. A number of observational data sets were combined for this purpose: SDSS data from Bell et al. (2003) and Baldry et al. (2004) at $z = 0$ and ULTRAVISTA (Ilbert et al. 2013; Muzzin et al. 2013) and ZFOURGE (Tomczak et al. 2014) at higher redshifts.

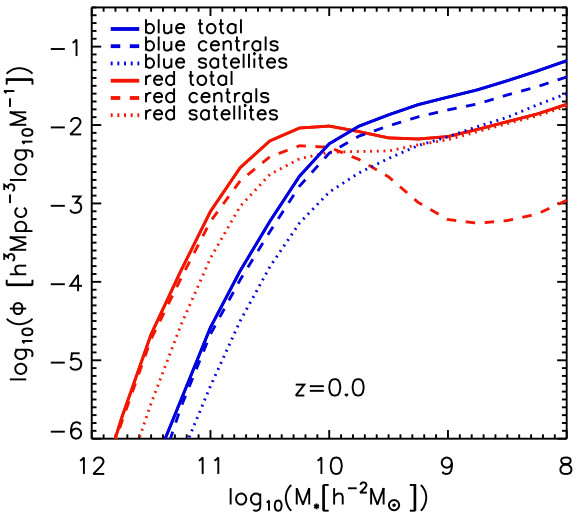

Figure 6. Low-redshift stellar mass functions for star-forming (blue lines) and passive (red lines) galaxies. Dashed lines give the functions for central galaxies and dotted lines the functions for satellites. The sums for each colour type are indicated by solid lines.

passive. Almost all passive, low-mass galaxies are satellites, but these red dwarfs are outnumbered by similar mass satellites which are still forming stars. Only for stellar masses above $5 \times 10^9 \mathrm{M}_\odot$ do red satellites outnumber blue ones. This reflects the fact that most low-mass satellites are in haloes where ram-pressure stripping is no longer effective.

Most previous semi-analytic models have overpredicted the passive fraction for low-mass galaxies (e.g. Bower et al. 2006; Croton et al. 2006; Weinmann et al. 2006a, 2012; Henriques et al. 2008; Henriques & Thomas 2010; Guo et al. 2011; Hirschmann et al. 2014). Our new reincorporation model ensures that central galaxies with $8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$ accrete gas and continue to form stars at late times, but this alone is not sufficient to reproduce the observations (see the dashed red lines in the left-hand panels of Fig 5). The observations require a significant fraction of low-mass satellites also to remain blue. Satellites in our model do not accrete primordial gas or wind ejecta and can only remain blue if the gas they already possess at infall remains available for star formation over a long period. This forces us to suppress ram-pressure stripping and to extend star formation in the absence of new accretion, changes which are similar to those introduced by Font et al. (2008) into their own galaxy formation model for similar reasons. Note that tidal stripping of material (gas, stars and dark matter) is still present in our new model. As we will show in Paper IV our new assumptions not only reproduce the observed abundances of galaxies as a function of stellar mass, star formation rate and redshift, but also the environmental dependences and quenching time-scales inferred from observations (Peng et al. 2010, 2012; Wetzel et al. 2012; Wang et al. 2014).

The new AGN feedback implementation is needed to make sure that galaxies with $M_\star \sim 10^{10.5}\,\mathrm{M}_\odot$ are predominantly passive at $z < 1$ despite late-time accretion of both primordial material and wind ejecta. (Compare the dashed and solid red lines in the left-hand panel of Fig 5.) For more massive galaxies, black hole and halo masses become sufficiently large to quench star formation even earlier.

5 FURTHER CONSEQUENCES OF DOWNSIZING

In the previous section, we focused on how and how well our new model adapts in order to fit the observational data which we use directly as constraints, namely the abundance and passive fraction of galaxies as functions of stellar mass from $z = 3$ to 0. Here, we analyse related quantities which were not used in our MCMC sampling in order to clarify the physical implications of the observed phenomenology. In particular: we study stellar mass functions separated explicitly into active and passive systems, together with the evolution of the implied cumulative mass and number densities; we compare observed and model distributions of colour, sSFR and luminosity-weighted stellar age for low-redshift galaxies; we analyse the evolution of the 'star-forming main sequence' of galaxies; and we look at the evolution of the mean cosmic star formation rate density (the Lilly–Madau diagram). A detailed comparison of observations with predicted black hole masses, cold-gas fractions, metallicities, morphologies and AGN feedback rates can be found in the Supplementary Material. Comparisons of clustering predictions with observations will be presented in Paper IV.

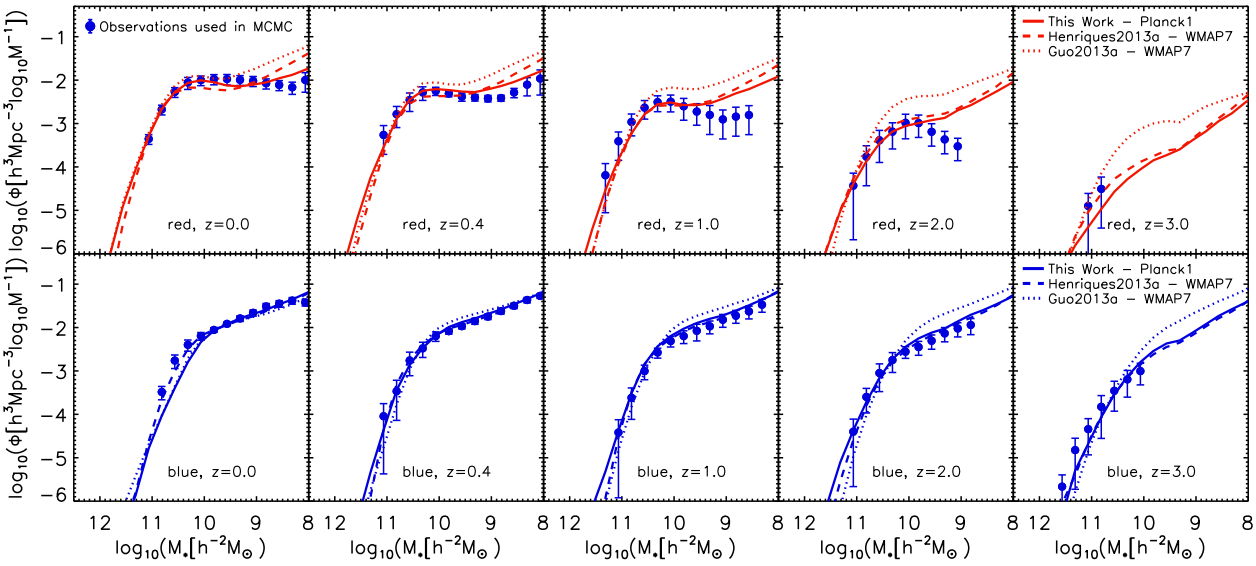

Figure 7. The evolution of the stellar mass functions of passive (top) and actively star forming (bottom) galaxies from $z = 3$ to 0. Theoretical predictions from Guo et al. (2013, the dotted lines), from Henriques et al. (2013, the dashed lines) and from our new model (the solid lines) are compared to our representation of the available observational data (blue symbols). As discussed in Appendix A, these include SDSS data at $z = 0$ (Bell et al. 2003; Baldry et al. 2004) and ZFOURGE (Tomczak et al. 2014) and ULTRAVISTA (Ilbert et al. 2013; Muzzin et al. 2013) data at higher redshifts. The individual observations are shown in Fig. A2. As in Fig. 2, the theoretical predictions have been convolved with a Gaussian with scatter that increases with redshift as suggested by Ilbert et al. (2013) in order to represent uncertainties in the observational stellar mass estimates. Estimates originally derived using Bruzual & Charlot (2003) stellar populations have been converted to Maraston (2005) populations by applying a correction of $\Delta M_\star = -0.14$ (Domínguez Sánchez et al. 2011).

5.1 Abundance evolution for active and passive galaxies

In Fig. 7, we separate the stellar mass functions already plotted in Fig. 2 into passive ('red', upper panels) and actively star forming ('blue', lower panels) systems using the colour cuts outlined in Section 4.2. The blue symbols are our representation of the results from several recent observational surveys, as detailed in Appendix A, and were used to estimate the passive fractions shown as blue symbols in Fig. 5. They are compared with model predictions from Guo et al. (2013, dotted lines), from Henriques et al. (2013, dashed lines) and from the model of this paper (solid lines). In this figure, the differences between our new model and Henriques et al. (2013) appear significant only at low redshift and are relatively small. Both models clearly represent the observations better than Guo et al. (2013), predicting fewer low-mass galaxies at early times and fewer low-mass, passive galaxies at all times. At low redshift, our new model has fewer of these $8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$ passive galaxies and more passive giants ($\log M_\star/\mathrm{M}_\odot \geq 11.0$) than Henriques et al. (2013) and is in reasonable agreement with the observations, although all models appear to predict more passive dwarfs than are observed at redshifts 1 and 2. The effect of our new AGN feedback is evident in the more rapid build-up of passive galaxies around the knee of the stellar mass function at late times.

Both in the observations and in our new model the stellar mass function of star-forming galaxies evolves very little between $z = 2$ and 0 and has a steep low-mass slope, whereas the mass function of passive galaxies has a much shallower low-mass slope and grows in amplitude by an order of magnitude while keeping the characteristic stellar mass at its knee almost constant. This behaviour has been noted previously (e.g. Bundy et al. 2006; Bell et al. 2007; Ilbert et al. 2010) in particular by Peng et al. (2010) who used it as the basis of a simple, toy model for galaxy formation. It implies that the growth of galaxies through star formation and mergers is being balanced by quenching processes that move galaxies from the active to the passive population. As noted by Peng et al. (2010), the constancy of the characteristic mass of passive galaxies implies that quenching typically occurs when galaxies grow to a well-defined stellar mass which is independent of time. In our physical model, this characteristic stellar mass is the minimum value for which feedback from the central supermassive black hole is able to offset cooling and accretion from the hot gas halo. As the passive galaxy stellar mass function grows in amplitude relative to that of active galaxies, the difference in shape between the two functions means that passive galaxies first start to dominate the population at high mass, and that the cross-over between active and passive domination moves steadily to lower stellar mass at later times.

A somewhat different view of this behaviour can be seen in Fig. 8. The upper panel shows comoving stellar mass density as a function of redshift for the galaxy population as a whole (black symbols and curve) and separated into the contributions from passive and active systems (red and blue symbols and curves, respectively). Symbols and shaded regions are obtained from our representation of the observed stellar mass functions and are integrated down to the observational completeness limits in each case. Predictions from our new model are shown integrated over all masses (dashed curves) and down to the same redshift- and colour-dependent completeness limits as the observation (solid curves). The stellar mass density in active systems is independent of redshift, once the variation in the completeness limit is accounted for, whereas the stellar mass density in passive systems increases by more than an order of magnitude to become dominant at $z = 0$. The lower panel of Fig. 8 shows cumulative comoving number densities of active and passive galaxies above three stellar mass thresholds as a function of redshift. Above the highest threshold, passive and active galaxies are already equally abundant at $z \sim 2$. Above the intermediate threshold, they are equally abundant at $z \sim 1$, while above the lowest threshold they only become equally abundant around $z \sim 0$. In both panels,

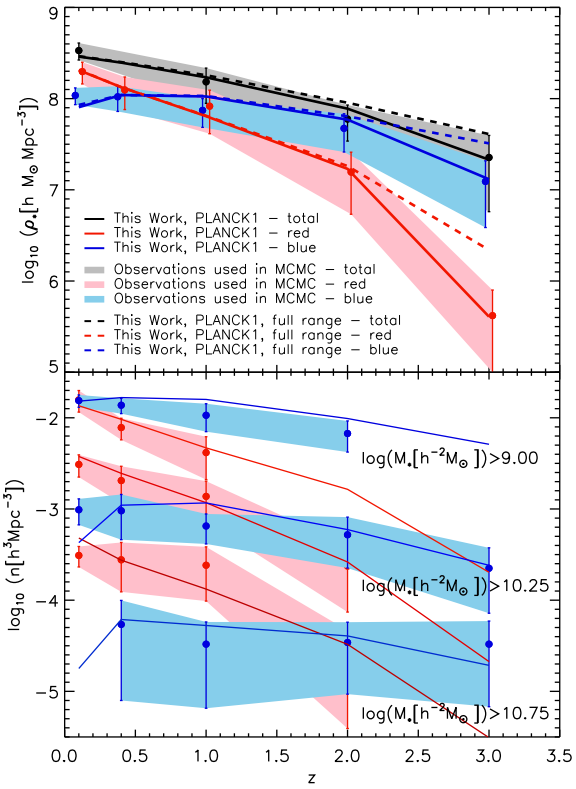

Figure 8. Upper panel: evolution of the comoving stellar mass density in passive (red) and active (blue) galaxies and in the full population (black). Symbols with error bars and the associated shaded regions are obtained from our compilation of observational results by integrating above the adopted completeness limit. Solid lines are predictions from the model of this paper integrated over all masses (dashed lines) and above the same completeness limits as the observations (solid lines). Lower panel: cumulative comoving number densities of active and passive galaxies above three stellar mass thresholds. Again, symbols and the associated shaded regions give observational results, while the solid lines are predictions from the model of this paper.

our new model is in qualitative agreement with the observational trends and appears to agree quantitatively within the considerable uncertainties.

5.2 Colours, sSFRs and ages of low-z galaxies

At low redshift the distributions of colour, sSFR and stellar age can be measured robustly for galaxies over a very wide mass range. In Fig. 9, we show $u-i$ colour distributions, r-band luminosity-weighted age distributions and sSFR distributions for SDSS galaxies split into eight disjoint bins spanning four orders of magnitude in stellar mass, $8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 12.0$. The observational data are shown as solid black histograms. sSFR estimates were derived as in Brinchmann et al. (2004), including the later corrections of Salim et al. (2007). In order to match the observed distribution of unclassified galaxies with no emission lines, that were assigned a value of SFR based on SED fitting, model galaxies with $\log \leq -12$ yr^{-1} have been assigned a random Gaussian value centred at $\log = -0.3\log(M_\star) - 8.6$ and with dispersion 0.5. Stellar masses are taken from Kauffmann et al. (2003) and stellar ages were derived as in Gallazzi et al. (2005). All observational data were downloaded from the MPA-JHU catalogue[4] and a $1/V_{\rm max}$ factor was applied when accumulating histograms in order to produce volume-limited statistics. As in previous plots, these observational data are compared with predictions from our new model, from Henriques et al. (2013) and from Guo et al. (2013).

For masses above $10^{10}\,\mathrm{M}_\odot$ there are clear differences between the three models. The delayed reincorporation of wind ejecta introduced by Henriques et al. (2013) produces an increase in the number of star-forming galaxies which is visible in the three left-most lower plots as a peak around 1–2 Gyr in the age distributions and around 10^{-10} yr^{-1} in the sSFR distributions. These peaks are less pronounced in the earlier model of Guo et al. (2013). As a result of this shift, Henriques et al. (2013) underpredicted the fraction of passive galaxies at high mass (see Fig. 5). Our updated treatment of feedback from radio mode AGN corrects this problem (although not completely) by quenching galaxies more strongly at later times. Its effects are evident in the reduction in the fraction of objects in these peaks in the new model. At high mass, galaxy colours are significantly redder in the two more recent models than in Guo et al. (2013) because of the change in population synthesis model and the larger value adopted for the nucleosynthetic yield. Discrepancies between model and observation nevertheless remain, particularly at intermediate masses, suggesting that our treatment of the transition between the low- and high-mass regimes, where the properties of galaxies change dramatically, still needs to be improved. In this range, the model distributions are much more clearly bimodal than those of SDSS galaxies. Another important factor is the treatment of dust which might also significantly contribute to this behaviour.

For low-mass galaxies ($8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$), the improved agreement of our new model with the SDSS data is quite clear. The peaks at red colours, old ages and low sSFR which were present to varying degrees in both the Henriques et al. (2013) and Guo et al. (2013) models are now strongly suppressed. Most dwarf galaxies are indeed blue, young and star forming, as observed, with only a small fraction of red/passive galaxies remaining. This reflects both the later reincorporation of ejecta, which ensures that almost all central galaxies are blue, and our altered assumptions about ram-pressure stripping and star formation thresholds, which allow a majority of low-mass satellite galaxies to remain blue. Although some disagreements of detail remain, the new model is in reasonable qualitative agreement with the inferred star formation rates of local galaxies over the full observed stellar mass range.

5.3 Evolution of the star-forming main sequence and the cosmic star formation rate density

In recent years, deep surveys of the galaxy population have identified a relation between star formation rate and stellar mass which holds for the bulk of star-forming systems. This 'main sequence' relation is close to a direct proportionality (e.g. Peng et al. 2010) and has relatively modest scatter at any given redshift, but its amplitude evolves strongly with time. The population of passive objects falls below this main sequence and has long been studied in detail at low redshift (e.g. Baldry et al. 2004; Brinchmann et al. 2004; Kauffmann et al. 2004). Recent increases in sensitivity have made similar analyses possible at high redshift, where these passive galaxies are the ones counted in the red' stellar mass functions of Fig. 7 which show them to be subdominant by number, even at high mass, for $z \geq 2$ (see also Fig. 8).

[4] http://www.mpa-garching.mpg.de/SDSS/

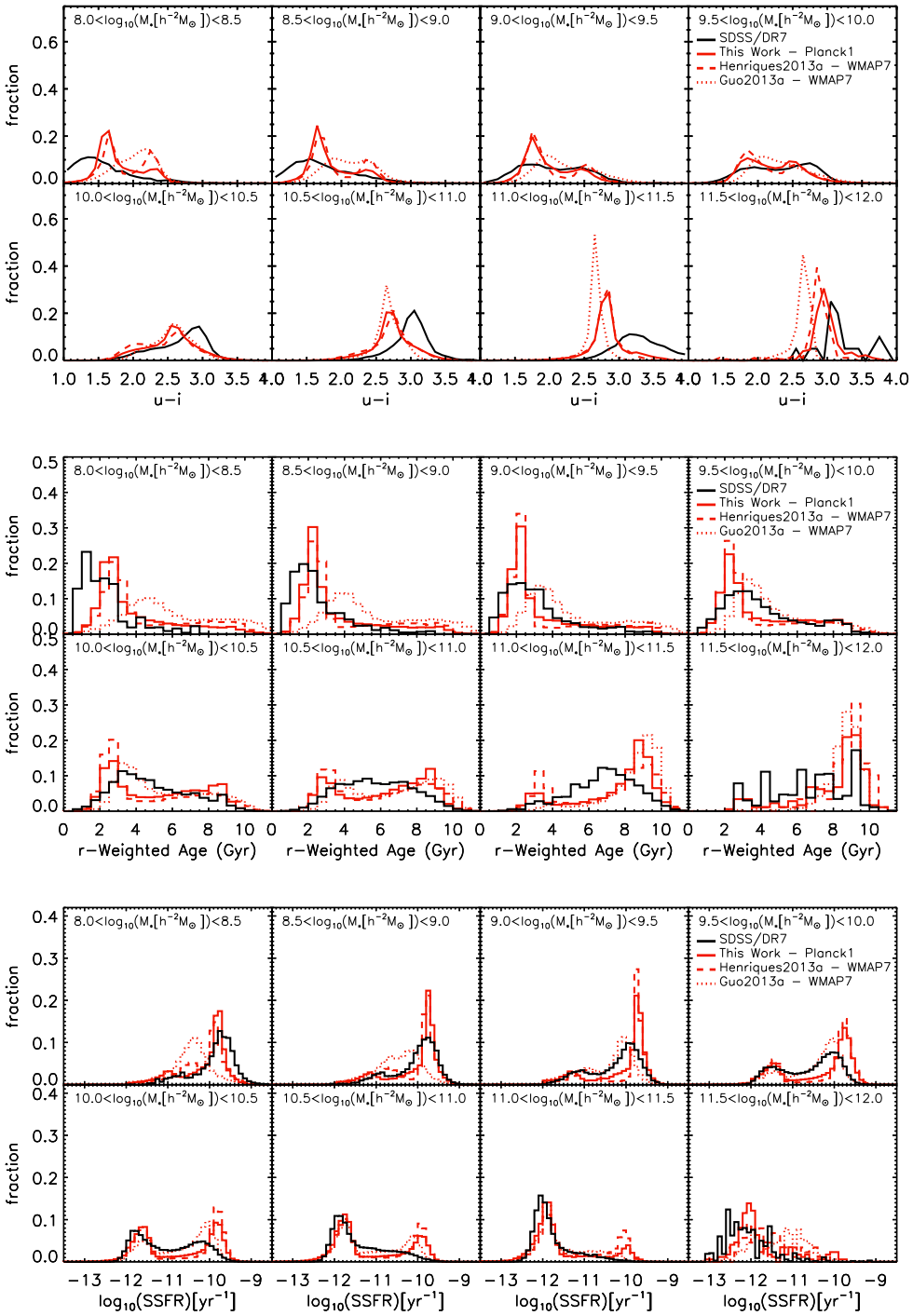

Figure 9. $u - i$ colour (top panels), r-band weighted age (middle panels) and sSFR (bottom panels) distributions at $z = 0.1$ in eight disjoint stellar mass bins. Observational estimates from the SDSS are shown as black solid histograms and are compared with predictions from our new model (solid red histograms), from Henriques et al. (2013, dashed red histograms) and from Guo et al. (2013, dotted red histograms). sSFR estimates for SDSS come from Brinchmann et al. (2004) and Salim et al. (2007), and luminosity-weighted stellar age estimates from Gallazzi et al. (2005).

Fig. 10 compares the redshift evolution of the relation between stellar mass and star formation rate in our new model (the grey-scale) with different observational determinations of the ridge-line of the star-forming main sequence (the solid lines and red symbols). The observed and model relations have similar slopes and evolve similarly with redshift. Both decrease in amplitude by well over an order of magnitude between $z = 2$ and 0. In addition, the scatter in the low-redshift relation is also quite similar in the model and in the data (see Fig. 9). Small discrepancies between theory and observations remain. These seem consistent with the calibration uncertainties of different observational star-formation rate indicators (see Kennicutt & Evans 2012 for a review on the subject) and with the detailed comparison of star formation rates in different models recently performed by Mitchell et al. (2014).

The cosmic star formation rate density is defined as the volume averaged sum of all ongoing star formation at any given time. Its

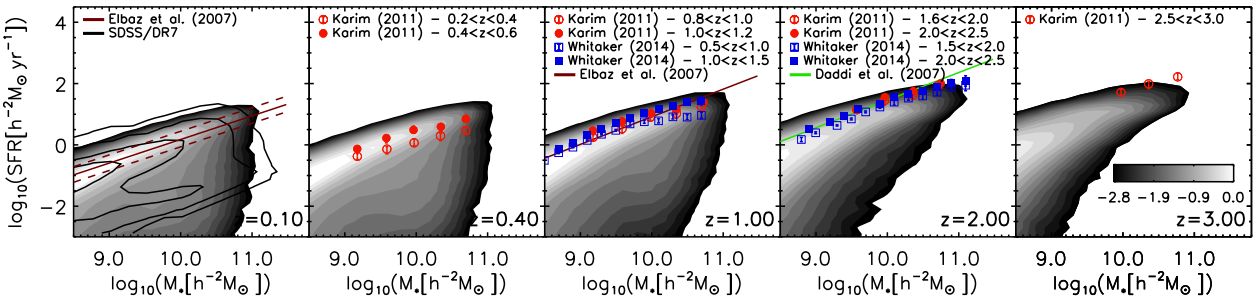

Figure 10. Star formation rate plotted against stellar mass from high to low redshift (right to left). The distributions predicted by the model of this paper are shown in grey-scale and are compared with symbols and straight lines representing the ridge-line of the observed 'star-forming main sequence' as inferred in recent studies that are indicated individually in each panel. At $z = 0.1$, we show the distribution inferred from SDSS with solid logarithmically spaced contours. The grey-scale contours represent the normalized number density of galaxies in logarithmically spaced bins (from light grey high-density to dark grey low-density contours).

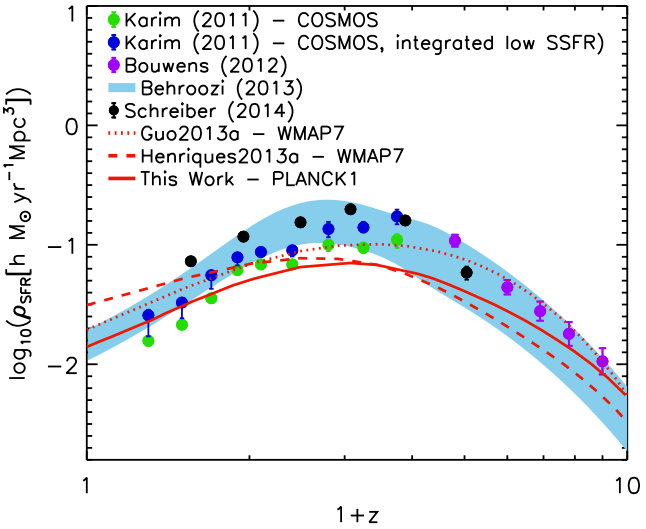

Figure 11. The evolution with redshift of the comoving density of cosmic star formation. The new model (solid red line), that of Henriques et al. (2013, dashed red line) and that of Guo et al. (2013, dotted red line) are compared with observational data from Karim et al. (2011), Bouwens et al. (2012), Schreiber et al. (2015) and Behroozi et al. (2013a).

integral from very early times down to redshift z should be equal to the integral of the mass-weighted stellar mass function at z, once mass-loss during stellar evolution is accounted for. In practice, however, the estimation of stellar mass and star formation rate densities from observational data is subject to large uncertainties and deviations between observational determinations of these two quantities are to be expected (Hopkins & Beacom 2006; Wilkins, Trentham & Hopkins 2008). As a result, a model that correctly predicts the observed evolution of the stellar mass function may not match observational estimates of the evolution of the star formation rate density.

In Fig. 11, we compare model predictions to observational estimates of the star formation rate density from $z = 9$ to 0. As in previous figures, we show results for the new model of this paper as a solid red line, results for the model of Henriques et al. (2013) as a dashed red line, and results for the Guo et al. (2013) model as a dotted red line. The observations are taken from COSMOS (Karim et al. 2011), the Bouwens et al. (2012) sample of Lyman-break galaxies, combined Herschel and *HST H* band-selected catalogues (Schreiber et al. 2015) and the Behroozi et al. (2013a) compilation. The observed rate has a broad peak at relatively low redshift ($z \sim 2$ to 3) and declines significantly by $z = 0$ but also to higher redshift. The prediction of these general features can be considered as one of the first significant successes of semi-analytic modelling of galaxy formation in a CDM universe (White 1989). Our new model matches the overall shape of the observed relation reasonably well although it is not peaked enough at $z = 2$. It seems that, despite fully matching the most recent observations of the stellar mass function from $z = 3$ to 0, we predict a milder decrease in the integrated star formation rate density than observed. There is thus some tension between the observational determination of these two quantities (Whitaker et al. 2014; Leja et al. 2015). Similar discrepancies were found by Furlong et al. (2014) when looking at the Schaye et al. (2015) numerical simulations.

We note that our modified treatment of radio mode feedback is responsible for the substantially larger drop in the star formation rate density at $z < 2$ than in the Henriques et al. (2013) model. The change from *WMAP7* to *Planck* cosmology results in higher halo accretion rates at early times and higher star formation rate densities at $z > 2$.

As pointed out by Schaye et al. (2010) and also seen in the MCMC analysis of Henriques et al. (2013), the high-redshift star formation rate density is mostly determined by the accretion of primordial material on to haloes with virial temperatures for which cooling is efficient. Below $z = 2$, the slowing of the cosmological accretion rate combines with a lengthening of characteristic cooling times to produce a global decrease in star formation rates. In addition, at later times AGN feedback and environmental quenching mechanisms also contribute to the decrease in the integrated star formation density, moving galaxies from the main star-forming sequence into the passive population.

6 SUMMARY AND CONCLUSIONS

We have updated the cosmological parameters underlying our galaxy formation model to the values preferred by the first analysis of *Planck* CMB data, while modifying our treatment of baryonic processes to address two major problems identified in earlier modelling, namely the overprediction of the abundance of low-mass galaxies ($8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 9.5$) at redshifts $z \geq 1$ and the overly large passive fraction predicted among low-redshift dwarfs. We use recent observational estimates of the abundances and passive fractions of galaxies over the stellar mass range $8.0 \leq \log M_\star/\mathrm{M}_\odot \leq 12.0$ and the redshift range $0 \leq z \leq 3$ as constraints on our modelling, using MCMC procedures to identify

the thresholds, scaling exponents and efficiencies needed for our treatment of baryonic processes to match the observations.

Relative to the most recent of our previous publicly-released models (Guo et al. 2011, 2013) matching these observations required us to delay the return of material ejected in galactic winds (as in Henriques et al. 2013), to weaken ram-pressure stripping in low-mass haloes with $\log M_{200}/\mathrm{M}_{\odot} < 14.0$ (as advocated by Font et al. 2008, for their own galaxy formation models), to lower the gas surface density threshold for star formation, and to make radio mode feedback from AGN more efficient at late times. With these changes, our new model reproduces our fiducial observations well over their full stellar mass and redshift ranges. In particular, it matches both the observed abundance of low-mass galaxies at $z \geq 1$ and the observed sharp, low-redshift transition between predominantly star-forming systems at low mass, $8.0 \leq \log M_{\star}/\mathrm{M}_{\odot} \leq 9.5$, and predominantly passive galaxies at high mass, $\log M_{\star}/\mathrm{M}_{\odot} > 10.5$. For low-redshift galaxies, the detailed distributions of colour, sSFR, and luminosity-weighted stellar age are matched reasonably well across the entire stellar mass range, $8.0 \leq \log M_{\star}/\mathrm{M}_{\odot} \leq 12.0$. In addition, the evolution of the mean cosmic star formation rate density over the range $0 < z < 9$ is reasonably well reproduced, once possible calibration uncertainties are allowed for.

Our new model embeds simple but plausible representations of the physical processes known to influence galaxy formation and evolution in the structure formation framework of the concordance ΛCDM model, yet it behaves in a very similar way to the simple toy model which Peng et al. (2010) introduced to interpret the observed evolution of stellar mass functions split into star forming and passive systems. At each redshift, there is a well-defined star-forming main sequence along which sSFR varies only weakly. The stellar mass function of star-forming galaxies has a steep low-mass slope and evolves very little with redshift, whereas that of passive galaxies has a much flatter low-mass slope and grows strongly in amplitude, but weakly in characteristic mass, with decreasing redshift. As a result, passive galaxies first come to dominate the population at high mass ($\log M_{\star}/\mathrm{M}_{\odot} \sim 11.3$ at $z \sim 2$) and the transition between active and passive domination shifts to progressively lower stellar mass at later times, dropping to $\log M_{\star}/\mathrm{M}_{\odot} \sim 10.0$ by $z = 0$. Peng et al. (2010) noted that fitting the data with their toy model required quenching of star formation to occur near a characteristic stellar mass which is independent of redshift. In our physical model, this characteristic stellar mass turns out to be the minimum value for which feedback from the central supermassive black hole is able to offset cooling and accretion from the hot gas halo.

In order to match the low passive fraction measured in low-redshift dwarf galaxies, we had to reduce environmental effects on dwarf satellites so that at least half of them are still star forming by the present day. This has an impact, of course, on the clustering of galaxies as a function of star formation activity, an issue which we will address in some detail in Paper IV, where we will show that the updates to our modelling also substantially improve the extent to which it matches observations of such 'environmental quenching'. Although many aspects of our baryonic modelling remain crude, and there are still some quantitative discrepancies with observations, we believe that the new model of this paper not only updates that of Guo et al. (2011) to the currently preferred *Planck* cosmology, but also cures the principal discrepancies with observations which were discovered in the earlier work without significantly damaging any of its successes. We therefore anticipate that publicly released data catalogues for the new model will be of considerable use for interpreting a broad variety of observations of the evolution and clustering of the galaxy population.

ACKNOWLEDGEMENTS

This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, and STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. The work of BH, SW and GL was supported by Advanced Grant 246797 'GALFORMOD' from the European Research Council. PT was supported by the Science and Technology Facilities Council [grant number ST/I000652/1]. GQ acknowledges support from the National basic research programme of China (973 programme under grant no. 2009CB24901), the Young Researcher Grant of National Astronomical Observatories, CAS, the NSFC grants programme (No. 11143005), as well as the Partner Group programme of the Max Planck Society. VS acknowledges support by the Deutsche Forschungsgemeinschaft through Transregio 33, 'The Dark Universe'. The authors thank Ivan Baldry, Olivier Ilbert, Alexander Karim, Cheng Li, Danilo Marchesini and Adam Muzzin for providing their observational data, Eric Bell, Jarle Brinchmann, Scott Clay, Guinevere Kauffmann, Bryan Terrazas, Loic Le Tiran, Jun Toshikawa, Stephen Wilkins and Rob Yates for useful discussions and Claudia Maraston, Gustavo Bruzual and Stephan Charlot for providing their stellar populations synthesis models.

REFERENCES

Angulo R. E., Hilbert S., 2015, MNRAS, 448, 364
Angulo R. E., White S. D. M., 2010, MNRAS, 405, 143
Arnouts S. et al., 2007, A&A, 476, 137
Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, ApJ, 600, 681
Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945
Baldry I. K. et al., 2012, MNRAS, 421, 621
Behroozi P. S., Wechsler R. H., Conroy C., 2013a, ApJ, 770, 57
Behroozi P. S., Marchesini D., Wechsler R. H., Muzzin A., Papovich C., Stefanon M., 2013b, ApJ, 777, L10
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS, 149, 289
Bell E. F., Zheng X. Z., Papovich C., Borch A., Wolf C., Meisenheimer K., 2007, ApJ, 663, 834
Benson A. J., 2012, 17, 175
Benson A. J., 2014, MNRAS, 444, 2599
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
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38
Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846
Bouwens R. J. et al., 2012, ApJ, 754, 83
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
Bower R. G., Vernon I., Goldstein M., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2010, MNRAS, 407, 2017
Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150
Brammer G. B. et al., 2011, ApJ, 739, 24
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Bundy K. et al., 2006, ApJ, 651, 120
Chabrier G., 2003, PASP, 115, 763
Charlot S., Bruzual G., 2007, in press

Clay S. J., Thomas P. A., Wilkins S. M., Henriques B. M. B., 2015, MNRAS, 451, 2692
Cole S., 1991, ApJ, 367, 45
Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781
Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486
Crain R. A. et al., 2015, preprint (arXiv:1501.01311)
Croton D. J. et al., 2006, MNRAS, 365, 11
De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499
Domínguez Sánchez H. et al., 2011, MNRAS, 417, 900
Drory N. et al., 2009, ApJ, 707, 1595
Faber S. M. et al., 2007, ApJ, 665, 265
Firmani C., Avila-Reese V., Rodríguez-Puebla A., 2010, MNRAS, 404, 1100
Font A. S. et al., 2008, MNRAS, 389, 1619
Fontana A. et al., 2006, A&AS, 459, 745
Fontanot F., Somerville R. S., Silva L., Monaco P., Skibba R., 2009, MNRAS, 392, 553
Fontanot F., Springel V., Angulo R. E., Henriques B., 2012, MNRAS, 426, 2335
Forcada-Miro M. I., White S. D. M., 1997, preprint (arXiv:astro-ph/9712204)
Fu J., Kauffmann G., Li C., Guo Q., 2012, MNRAS, 424, 2701
Fu J. et al., 2013, MNRAS, 434, 1531
Furlong M. et al., 2014, preprint (arXiv:1410.3485)
Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41
Genel S. et al., 2014, preprint (arXiv:1405.3749)
Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2004, ApJ, 600, 580
Guo Q. et al., 2011, MNRAS, 413, 101
Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, MNRAS, p. 135
Hatton S., Devriendt J. E. G., Ninin S., Bouchet F. R., Guiderdoni B., Vibert D., 2003, MNRAS, 343, 75
Henriques B. M. B., Thomas P. A., 2010, MNRAS, 403, 768
Henriques B. M., Bertone S., Thomas P. A., 2008, MNRAS, 383, 1649
Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, MNRAS, 396, 535
Henriques B., Maraston C., Monaco P., Fontanot F., Menci N., De Lucia G., Tonini C., 2011, MNRAS, 415, 3571
Henriques B. M. B., White S. D. M., Lemson G., Thomas P. A., Guo Q., Marleau G.-D., Overzier R. A., 2012, MNRAS, 421, 2904
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, MNRAS, 431, 3373
Hirschmann M., De Lucia G., Wilman D., Weinmann S., Iovino A., Cucciati O., Zibetti S., Villalobos Á., 2014, MNRAS, 444, 2938
Hopkins A. M., Beacom J. F., 2006, ApJ, 651, 142
Ilbert O. et al., 2010, ApJ, 709, 644
Ilbert O. et al., 2013, A&AS, 556, A55
Kampakoglou M., Trotta R., Silk J., 2008, MNRAS, 384, 1414
Kang X., van den Bosch F. C., 2008, ApJ, 676, L101
Kang X., Jing Y. P., Mo H. J., Börner G., 2005, ApJ, 631, 21
Karim A. et al., 2011, ApJ, 730, 61
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Kauffmann G., Colberg J., Diaferio A., White S., 1999, MNRAS, 303, 188
Kauffmann G. et al., 2003, MNRAS, 341, 33
Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, MNRAS, 353, 713
Kauffmann G., Li C., Zhang W., Weinmann S., 2013, MNRAS, 430, 1447
Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531
Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Lacey C., Silk J., 1991, ApJ, 381, 14
Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011, MNRAS, 418, 1649
Leja J., van Dokkum P., Franx M., 2013, ApJ, 766, 33
Leja J., van Dokkum P., Franx M., Whitaker K., 2015, ApJ, 798, 115
Lemson G., Virgo Consortium, 2006, preprint (arXiv:astro-ph/0608019)
Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782
Li C., White S. D. M., 2009, MNRAS, 398, 2177
Lu Y., Mo H. J., Weinberg M. D., Katz N., 2011, MNRAS, 416, 1949
Lu Y., Mo H. J., Katz N., Weinberg M. D., 2012, MNRAS, 421, 1779
Lu Y. et al., 2014, ApJ, 795, 123
McCarthy I. G., Frenk C. S., Font A. S., Lacey C. G., Bower R. G., Mitchell N. L., Balogh M. L., Theuns T., 2008, MNRAS, 383, 593
Maraston C., 2005, MNRAS, 362, 799
Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, ApJ, 701, 1765
Marchesini D. et al., 2010, ApJ, 725, 1277
Mitchell P. D., Lacey C. G., Baugh C. M., Cole S., 2013, MNRAS, 435, 87
Mitchell P. D., Lacey C. G., Cole S., Baugh C. M., 2014, MNRAS, 444, 2637
Mutch S. J., Poole G. B., Croton D. J., 2013, MNRAS, 428, 2001
Muzzin A. et al., 2013, ApJ, 777, 18
Neistein E., Weinmann S. M., 2010, MNRAS, 405, 2717
Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577
Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, MNRAS, 406, 2325
Papovich C., Finkelstein S. L., Ferguson H. C., Lotz J. M., Giavalisco M., 2011, MNRAS, 412, 1123
Peng Y.-j. et al., 2010, ApJ, 721, 193
Peng Y.-j., Lilly S. J., Renzini A., Carollo M., 2012, ApJ, 757, 4
Pforr J., Maraston C., Tonini C., 2012, MNRAS, 422, 3285
Planck Collaboration XVI, 2014, A&A, 571, A16
Pozzetti L. et al., 2010, A&AS, 523, A13
Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
Ruiz A. N. et al., 2015, ApJ, 801, 139
Salim S. et al., 2007, ApJS, 173, 267
Schaye J. et al., 2010, MNRAS, 402, 1536
Schaye J. et al., 2015, MNRAS, 446, 521
Schreiber C. et al., 2015, A&A, 575, A74
Shamshiri S., Thomas P. A., Henriques B. M., Tojeiro R., Lemson G., Oliver S. J., Wilkins S., 2015, MNRAS, 451, 2681
Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V. et al., 2005, Nature, 435, 629
Sun M., 2012, New J. Phys., 14, 045004
Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, ApJ, 621, 673
Tomczak A. R. et al., 2014, ApJ, 783, 85
van Dokkum P. G. et al., 2010, ApJ, 709, 1018
Vogelsberger M. et al., 2014, Nature, 509, 177
Wang J., De Lucia G., Kitzbichler M. G., White S. D. M., 2008, MNRAS, 384, 1301
Wang L., Weinmann S. M., Neistein E., 2012, MNRAS
Wang W., Sales L. V., Henriques B. M. B., White S. D. M., 2014, MNRAS, 442, 1363
Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006a, MNRAS, 366, 2
Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., Croton D. J., Moore B., 2006b, MNRAS, 372, 1161
Weinmann S. M., Kauffmann G., von der Linden A., De Lucia G., 2010, MNRAS, 406, 2249
Weinmann S. M., Pasquali A., Oppenheimer B. D., Finlator K., Mendel J. T., Crain R. A., Macciò A. V., 2012, MNRAS, 426, 2797
Wetzel A. R., Tinker J. L., Conroy C., 2012, MNRAS, 424, 232
Whitaker K. E. et al., 2014, ApJ, 795, 104
White S. D. M., 1989, in Frenk C. S., Ellis R. S., Shanks T., Heavens A. R., Peacock J. A., eds, Proc. NATO ASIC Vol. 264, The Epoch of Galaxy Formation Observable Signatures of Young Galaxies. Kluwer, Dordrecht, p. 15
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White C. E., Somerville R. S., Ferguson H. C., 2015, ApJ, 799, 201
Wilkins S. M., Trentham N., Hopkins A. M., 2008, MNRAS, 385, 687
Yoshida N., Stoehr F., Springel V., White S. D. M., 2002, MNRAS, 335, 762

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

supplementary_material.zip
(http://mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stv705/-/DC1).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX A: COMPARING MODELS AND OBSERVATIONS

In this appendix, we define the figure of merit used to assess the level of agreement between our models and the observational data with which we constrain them. In addition, we show the individual observational data sets and describe how they are combined to give constraints with realistic uncertainty estimates which are suitable for MCMC exploration of the high-dimensional parameter space of our models. A more detailed description of our methods can be found in appendix 3 of Henriques et al. (2013).

A1 Figure of merit

Our figure of merit for each model is its 'likelihood', given the constraining observations. This is computed assuming individual data points to be independently and normally distributed around the model prediction with variance corresponding to the sum in quadrature of observational and theoretical uncertainties estimated as detailed below. The observational uncertainties are dominated by systematics rather than by sampling noise, so neither the Gaussian assumption nor the precise variance can be rigorously justified. In addition, both types of uncertainty are expected to be strongly correlated between data points. As a result, our figure of merit is not a true likelihood and our MCMC analysis should be interpreted as indicating acceptable regions of parameter space, rather than as defining posterior probability distributions in a rigorous Bayesian sense (see Benson 2014, for related discussion).

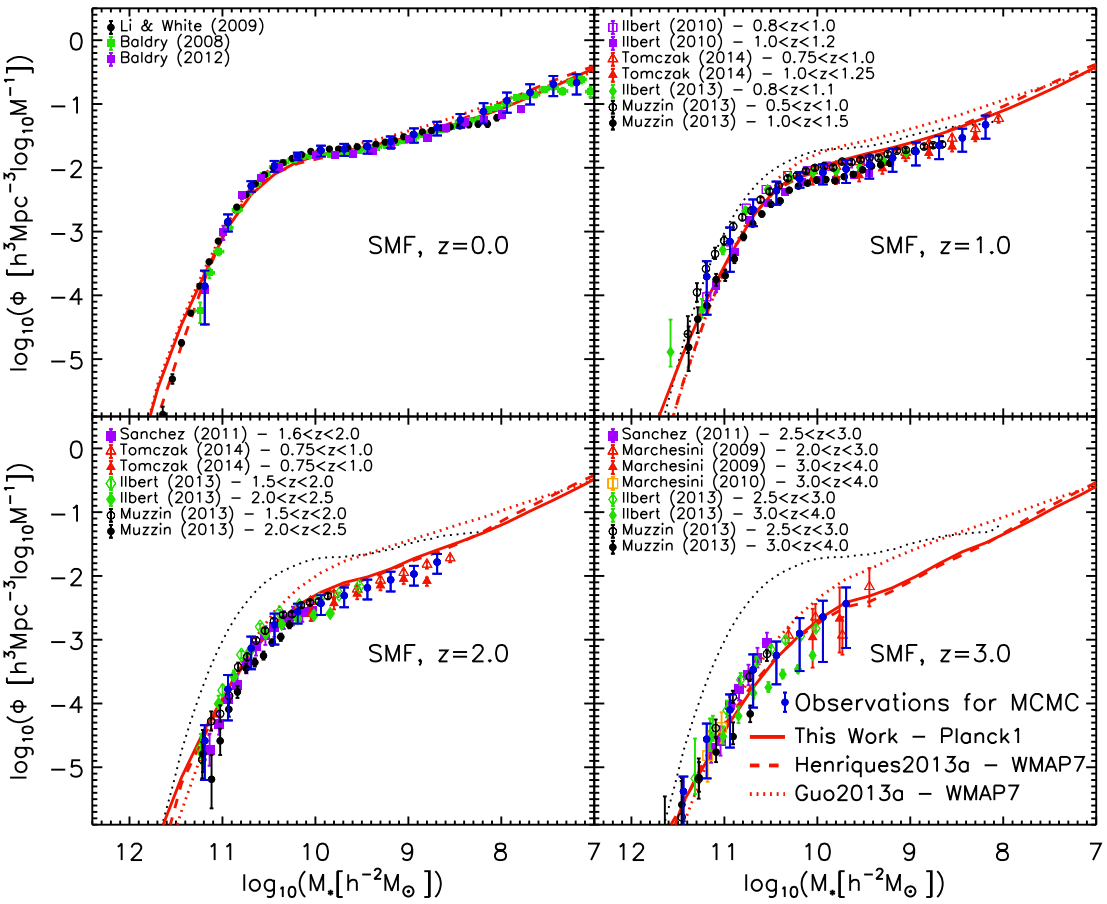

Figure A1. Evolution of the stellar mass function from $z = 3$ to 0 as in Fig. 2, except with the data points for the individual underlying surveys also shown. These surveys include SDSS (Baldry et al. 2008; Li & White 2009) and GAMA (Baldry et al. 2012) at $z \sim 0$ and Marchesini et al. (2009), Spitzer-COSMOS (Ilbert et al. 2010), NEWFIRM (Marchesini et al. 2010), COSMOS (Domínguez Sánchez et al. 2011), ULTRAVISTA (Ilbert et al. 2013; Muzzin et al. 2013) and ZFOURGE (Tomczak et al. 2014) at higher redshifts. All mass estimates at $z > 0$, except Domínguez Sánchez et al. (2011) and Muzzin et al. (2013) have been shifted by −0.14 dex to convert from Bruzual & Charlot (2003) to Maraston (2005) stellar populations (Domínguez Sánchez et al. 2011). The $z = 0$ results of Li & White (2009) are repeated at all redshifts as a black dotted line.

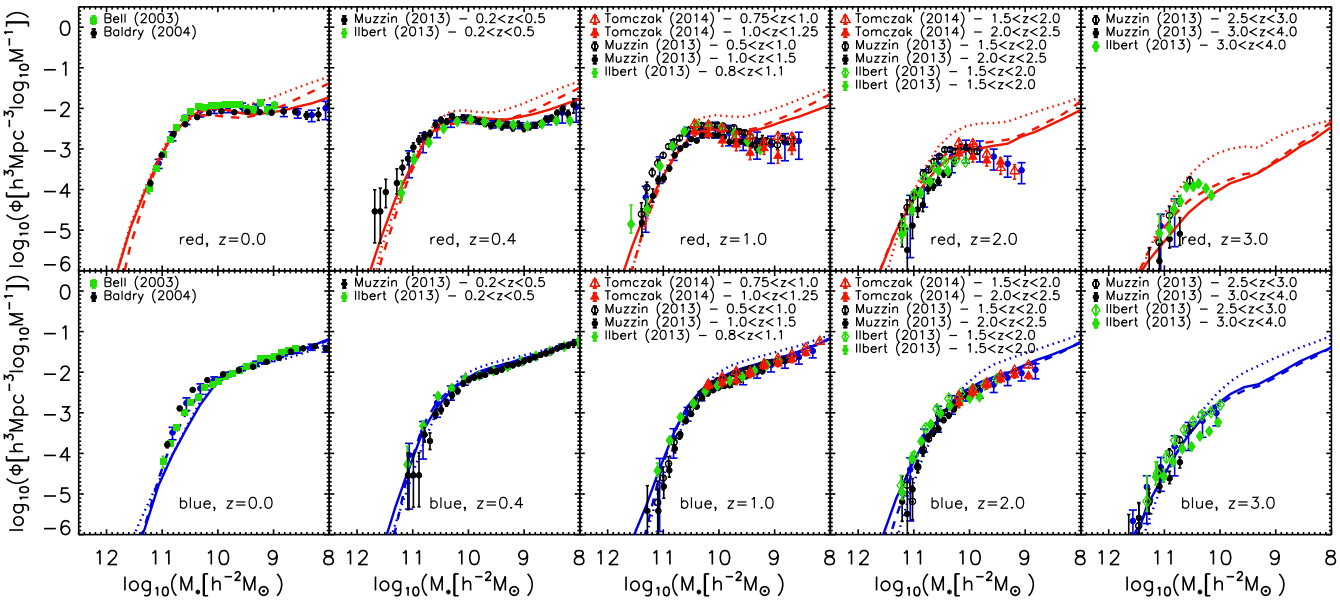

Figure A2. Evolution of the stellar mass function of red and blue galaxies from $z = 3$ to 0 as in Fig. 7, but with individual observational data sets shown. These include SDSS data from Bell et al. (2003) and Baldry et al. (2004) at $z = 0$ and ULTRAVISTA (Ilbert et al. 2013; Muzzin et al. 2013) and ZFOURGE (Tomczak et al. 2014) at higher redshifts. The Ilbert et al. (2013) and Tomczak et al. (2014) data have been shifted by −0.14 dex to convert from Bruzual & Charlot (2003) to Maraston (2005) stellar populations (Domínguez Sánchez et al. 2011). For Ilbert et al. (2013) data points this was done even for their lowest redshift bin ($0.2 < z < 0.5$) in order to ensure consistency with the Muzzin et al. (2013) data.

Our procedure is similar to that outlined in Henriques et al. (2013). We again use observed stellar mass functions at a series of redshifts ($z = 0$, 1, 2 and 3) to constrain galaxy abundances, but, rather than using *K*- and *B*-band luminosity functions to constrain the relative numbers of passive and actively star-forming galaxies, we here use direct estimates of the passive fraction as a function of stellar mass at $z = 0$, 0.4, 1, 2 and 3. These are obtained from stellar mass functions for galaxy samples split into active and passive subsets according to colour–colour plots like Fig. 4. This change makes our analysis less sensitive to the details of stellar population synthesis models and separates constraints on galaxy abundance more clearly from constraints on galaxy activity.

Given the above assumptions our figure of merit for each model (its 'likelihood') can be computed as

$$\mathcal{L} \propto \exp\left(-\chi^2/2\right), \tag{A1}$$

where χ^2 is given by

$$\chi^2 = \sum_{i,j} \frac{(\phi_{i,j} - \tilde{\phi}_{i,j})^2}{\Delta\phi_{i,j}^2 + \Delta\tilde{\phi}_{i,j}^2} + \sum_{i,j} \frac{(f_{i,j} - \tilde{f}_{i,j})^2}{\Delta f_{i,j}^2 + \Delta\tilde{f}_{i,j}^2}. \tag{A2}$$

The first sum on the r.h.s. of this equation involves the stellar mass functions ϕ (here defined as the logarithm of the number density of galaxies of the relevant mass and redshift), while the second involves the passive fractions f. In each case, the index i enumerates the redshifts used, while j enumerates the stellar mass bins at each redshift. Observational quantities and their adopted uncertainties are indicated by a tilde, while model predictions and their uncertainties have no tilde. The uncertainty in a quantity x is denoted by Δx.

We compute passive fractions as a function of stellar mass for all modern surveys for which the authors have explicitly estimated stellar mass functions separately for active and passive systems. As discussed in the next subsection, we combine these independent estimates to obtain the stellar mass functions, $\tilde{\phi}_B$ and $\tilde{\phi}_R$ and associated uncertainties shown in Fig. 7. These functions are then combined to give the passive fractions we use as constraints

$$\tilde{f}_{i,j} = \frac{\tilde{\phi}_{R;i,j}}{\tilde{\phi}_{R;i,j} + \tilde{\phi}_{B;i,j}}. \tag{A3}$$

The uncertainties in the passive fractions are straightforwardly obtained from those in the stellar mass functions by standard error propagation,

$$\Delta\tilde{f} = \frac{\tilde{\phi}_R\tilde{\phi}_B}{(\tilde{\phi}_R + \tilde{\phi}_B)^2}\sqrt{(\Delta\tilde{\phi}_R/\tilde{\phi}_R)^2 + (\Delta\tilde{\phi}_B/\tilde{\phi}_B)^2}, \tag{A4}$$

where the index pair i, j has been suppressed on all quantities for clarity.

Finally, we assume the theoretical uncertainty on the predicted passive fractions to be $\Delta f_{i,j} = 0.025$, based on the scatter in the passive fraction among the tree subsamples used in our MCMC analysis, and we neglect the theoretical uncertainty in the stellar mass functions, setting $\Delta\phi_{i,j} = 0$.

A2 Individual observational data sets

As in Henriques et al. (2009), Henriques & Thomas (2010) and Henriques et al. (2013), we combine multiple determinations of each stellar mass function, using the scatter among them to indicate the systematic uncertainties which appear in most cases to be larger than those purely due to count statistics. For each redshift range and for each stellar mass bin, we take a straight average of the different data sets and assume the 1σ uncertainty to be half of the maximum to minimum range. By not weighting the averages we attempt to recognize the fact that systematic errors can affect large and small surveys in similar ways. However, we emphasize

that this estimate of uncertainties is crude, and that in the presence of systematics any comparison between theory and observations is essentially qualitative. Formal levels of agreement should thus be treated with considerable caution.

Our adopted constraints are shown together with the individual data sets on which they are based in Fig. A1 for the overall stellar mass function and in Fig. A2 for the stellar mass functions of passive and actively star-forming galaxies. The constraints we use in our MCMC sampling are shown as blue dots with error bars, while other data points represent published observational estimates from the individual surveys. Theoretical predictions are shown as red lines.

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