The dark side of galaxy colour

Andrew P. Hearin[1]★ and Douglas F. Watson[2,3]†

[1]*Fermilab Center for Particle Astrophysics, Fermi National Accelerator Laboratory, Batavia, IL 60510-0500, USA*
[2]*Department of Astronomy & Astrophysics, The University of Chicago, Chicago, IL 60637, USA*
[3]*Kavli Institute for Cosmological Physics, 5640 South Ellis Avenue, The University of Chicago, Chicago, IL 60637, USA*

ABSTRACT

We present *age distribution matching*, a theoretical formalism for predicting how galaxies of luminosity L and colour C occupy dark matter haloes. Our model supposes that there are just two fundamental properties of a halo that determine the colour and brightness of the galaxy it hosts: the maximum circular velocity V_{max} and the redshift z_{starve} that correlates with the epoch at which the star formation in the galaxy ceases. The halo property z_{starve} is intended to encompass physical characteristics of halo mass assembly that may deprive the galaxy of its cold gas supply and, ultimately, quench its star formation. The new, defining feature of the model is that, at fixed luminosity, galaxy colour is in monotonic correspondence with z_{starve}, with the larger values of z_{starve} being assigned redder colours. We populate an *N*-body simulation with a mock galaxy catalogue based on age distribution matching and show that the resulting mock galaxy distribution accurately describes a variety of galaxy statistics. Our model suggests that halo and galaxy assembly are indeed correlated. We make publicly available our low-redshift, Sloan Digital Sky Survey $M_r < -19$ mock galaxy catalogue, and main progenitor histories of all $z = 0$ haloes, at http://logrus.uchicago.edu/~aphearin

Key words: galaxies: evolution – galaxies: haloes – cosmology: theory – dark matter – large-scale structure of Universe.

1 INTRODUCTION

Galaxies are not colour blind in how they congregate with one another. Redder galaxies (older with little active star formation) tend to occupy more dense environments, and conversely bluer galaxies (younger with active ongoing star formation) preferentially reside in underdense regions (Balogh et al. 1999; Blanton et al. 2005; Weinmann et al. 2006, 2009; Peng et al. 2010, 2012; Wetzel, Tinker & Conroy 2012; Carollo et al. 2012). The colour dependence of galaxy location within the cosmic web also manifests in measurements of two-point statistics (e.g. Norberg et al. 2002; Zehavi et al. 2002, 2005, 2011; Li et al. 2006; Yang et al. 2012; Mostek et al. 2013). Additionally, there is a clear bimodality in the distribution of galaxy colour and brightness, with distinct red and blue sequences in colour–magnitude space, along with a minority 'green valley' population (Blanton et al. 2003, 2005; Baldry et al. 2004; Wyder et al. 2007). This segregation between blue star-forming and red passively evolving galaxies has been shown to persist out to $z \sim 1$ (Bell et al. 2004; Cooper et al. 2006, 2012) and even as high as $z \sim 3$ (Whitaker et al. 2011). While these correlated trends in colour–magnitude and configuration space are observationally well established, our theoretical understanding of the physical processes driving these correlations remains incomplete.

In the standard Λ cold dark matter (ΛCDM) cosmological model, structures assemble hierarchically, wherein high-density regions condense and virialize, forming bound structures known as dark matter haloes. Haloes of sufficient mass are the natural sites for luminous galaxies to form (White & Rees 1978; Blumenthal et al. 1984), and the way galaxies connect to dark matter haloes provides the fundamental link between predictions of galaxy formation theory, such as colour, and the concordance ΛCDM model. Understanding this link requires detailed knowledge of the spatial distribution of galaxies and haloes, and towards this end large galaxy redshift surveys such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) have produced detailed three-dimensional maps of galaxies in the local Universe.

Extensive effort has been put forth to understand how, as a function of colour, galaxies map to dark matter haloes in a cosmological context, and there are two general approaches to constructing the galaxy–halo connection. First, there is the halo occupation distribution (HOD) technique [and the closely related conditional luminosity function (CLF) technique (e.g. Yang, Mo & van den Bosch 2003; van den Bosch et al. 2007)], which adopts parametric forms for the biased relation between galaxies and halo mass (e.g. Peacock & Smith 2000; Seljak 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2002; Cooray & Sheth 2002; Zheng, Coil &

★ E-mail: aphearin@fnal.gov
† NSF Astronomy, and Astrophysics Postdoctoral Fellow.

Zehavi 2007; Leauthaud et al. 2012; Watson et al. 2012a). The HOD describes this relation by specifying the probability distribution that a halo of mass M contains N galaxies, along with a prescription for the spatial distribution of galaxies within haloes. In this formalism, it is typically assumed that a single 'central' galaxy lives at the centre of each distinct halo, with additional 'satellite' galaxies tracing the dominant dark matter component. The most common implementations of this approach do not explicitly use any information about dark matter subhaloes (self-bound, dark matter structures orbiting within their host halo potential).

The second technique is known as abundance matching (Kravtsov et al. 2004; Tasitsiomi et al. 2004; Vale & Ostriker 2004, 2006; Conroy, Wechsler & Kravtsov 2006; Conroy & Wechsler 2009; Guo et al. 2010; Neistein et al. 2011a; Rodríguez-Puebla, Drory & Avila-Reese 2012; Simha et al. 2012; Watson, Berlind & Zentner 2012b; Hearin et al. 2013b; Kravtsov 2013; Reddick et al. 2013). This is a simple, yet remarkably powerful approach wherein by assuming a monotonic relation between the maximum circular velocity V_{max} (or halo mass) and luminosity L (or stellar mass), one can 'abundance match' to make the correspondence between a dark matter (sub)halo and the galaxy residing at its centre. In this case, galaxies are assigned to dark matter haloes with the most massive galaxies residing in the most massive haloes. This yields an implicit relationship between L and V_{max} such that, by construction, the luminosity function of mock galaxy catalogues is in exact agreement with the data.

The most accurate and precise theoretical descriptions of the galaxy colour–halo connection to date have been predominantly derived from modelling the colour-dependent galaxy two-point correlation function (2PCF) through the HOD formalism (e.g. Zehavi et al. 2005, 2011; Skibba & Sheth 2009; Tinker & Wetzel 2010; Krause et al. 2013) or semi-analytic models (SAMs; White & Frenk 1991; Kauffmann et al. 1999). However, abundance matching-type theories implemented in a cosmological N-body simulation are now being explored (Gerke et al. 2013; Masaki, Lin & Yoshida 2013). Alternatively, new phenomenological approaches that seek to simplify the complexities of SAMs are actively being pursued (Peng et al. 2010; Lilly et al. 2013; Mutch, Croton & Poole 2013).

Our intent in this paper is to try and uncover the simplest, physically motivated model for associating the colour of a galaxy with the mass accretion history of the galaxy's parent halo. To that end, we introduce a new theoretical formalism which we call *age distribution matching*. This model supposes that there are just two fundamental properties of a halo that determine the colour and brightness of the galaxy it hosts: (1) the maximum circular velocity of a test particle orbiting in its gravitational potential well (i.e. V_{max}) and (2) the redshift at which the galaxy will be starved of cold gas, which we refer to as z_{starve} throughout the paper. This halo property z_{starve} is intended to encompass characteristics of a halo's assembly history which should be associated with physical processes that can deprive a galaxy of a cold gas supply and, ultimately, quench star formation. The novel feature of the model is that, at fixed luminosity, galaxy colour is in monotonic correspondence with z_{starve}. By constructing a mock galaxy catalogue with both r-band magnitudes (luminosities) M_r and $(g - r)$ colours, we compare the predictions of our model to a number of observed galaxy statistics.

The rest of this paper is organized as follows. In Section 2, we discuss the data incorporated throughout this work and in Section 3 we discuss the simulation, halo catalogues and merger trees used in this study. In Section 4, we describe our algorithm for constructing our SDSS-based mock catalogue including a detailed description of our age distribution matching formalism. Results are presented

in Section 5 followed by a discussion in Section 6. Finally, in Section 7 we give a summary of our work. Throughout we assume a flat ΛCDM cosmological model with $\Omega_m = 0.27$ and all quoted magnitudes use $h = 1$.

2 DATA

We construct our low-redshift mock galaxy sample based on the galaxy catalogue described in Section 2.1. In particular, this is the sample that defines the distribution of colours at fixed luminosity, $P_{SDSS}(g - r|L)$, required in age distribution matching (see Section 4.1.2). In Section 2.2, we describe the projected 2PCF measurements that we use to compare with the clustering predictions of our mock sample. We use the galaxy group sample described in Section 2.3 to test our predictions for how galaxies arrange themselves into groups by colour.

2.1 SDSS galaxy sample

For our baseline galaxy sample, we use a volume-limited catalogue of galaxies constructed from the main galaxy sample of Data Release 7 (DR7 hereafter; Abazajian et al. 2009) of SDSS. This galaxy catalogue is an update of the Berlind et al. (2006) sample, which was based on SDSS Data Release 3, to which we refer the reader for details. Briefly, our volume-limited spectroscopic subsample ($V_{eff} = 5.8 \times 10^6 \, h^{-3} \, \text{Mpc}^3$) spans the redshift range $0.02 \leq z \leq 0.068$ with r-band absolute magnitude $M_r - 5 \log h < -19$. For M_r, we use Petrosian magnitude measurements. Throughout this paper, we refer to this as the 'Mr19' galaxy sample.

The term 'fibre collisions' refers to the cases when the angular separation between two or more galaxies is closer than the minimum separation (55 arcsec) permitted by the plugging mechanism of the optical fibres in the spectrometer (see Guo, Zehavi & Zheng 2012, and references therein). The treatment of fibre collisions in the DR7-based galaxy sample we use in this paper differs from that in the catalogue presented in Berlind et al. (2006). These differences were discussed in detail in appendix B of Hearin et al. (2013b), to which we refer the reader for details.

2.2 Clustering measurements

We compare our model predictions for the colour-dependent 2PCF to the SDSS measurements presented in Zehavi et al. (2011). Using the observed $(g - r)$ colours, they separate red and blue galaxy populations based on the following magnitude-dependent colour cut:

$$(g - r) = 0.21 - 0.03 M_r. \tag{1}$$

We use the luminosity-binned clustering data presented in their tables C9 and C10 to test our model 2PCF predictions for red and blue galaxies.

2.3 Galaxy group measurements

Galaxy groups are constructed from the Mr19 galaxy sample described in Section 2.1 according to the methods discussed in Berlind et al. (2006). Briefly, groups are identified via a redshift-space friends-of-friends algorithm that takes no account of member galaxy properties beyond their redshifts and positions on the sky. For each group, we define the central galaxy to be the brightest

non-fibre-collided member,[1] and the group's satellites to be the remaining non-fibre-collided members. The details of how fibre collisions influence central/satellite designations in this group sample were studied extensively in Hearin et al. (2013b, hereafter H13), and we follow their conventions here. We refer the reader to Berlind et al. (2006) for further details on the group-finding algorithm.

3 SIMULATION AND HALO CATALOGUES

Throughout this work, we use the high-resolution, collisionless, CDM *N*-body Bolshoi simulation (Klypin, Trujillo-Gomez & Primack 2011). The simulation has a volume of 250^3 h^{-3} Mpc^3 with 2048^3 particles and a ΛCDM cosmological model with $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.042$, $h = 0.7$, $\sigma_8 = 0.82$ and $n_s = 0.95$. Bolshoi has a mass resolution of 1.9×10^8 M$_\odot$ and force resolution of 1 h^{-1} kpc, and particles were tracked from $z = 80$ to 0. Bolshoi was run with the Adaptive Refinement Tree code (Kravtsov, Klypin & Khokhlov 1997; Gottloeber & Klypin 2008). The Bolshoi data are available at http://www.multidark.org and we refer the reader to Riebe et al. (2011) for additional information.

Haloes and subhaloes were identified with the phase-space temporal halo finder ROCKSTAR (Behroozi et al. 2013b,c), capable of resolving Bolshoi haloes and subhaloes down to $V_{max} \sim 55$ km s^{-1}. Halo masses were calculated using spherical overdensities according to the redshift-dependent virial overdensity formalism of Bryan & Norman (1998). We use the $z = 0$ ROCKSTAR halo catalogue and merger trees to model the galaxy colour–halo connection.

4 MOCK CATALOGUES

In this section, we describe our algorithm for constructing our mock sample of galaxies and galaxy groups. We apply age distribution matching, defined in Section 4.1, to the ROCKSTAR halo catalogues at $z = 0$ to construct a sample of mock galaxies with both *r*-band luminosities M_r and $(g - r)$ colours. After constructing the mock galaxy sample from the halo catalogue, we place the mock galaxies into redshift space and apply the same group finder to the mock data that was applied to the SDSS Mr19 galaxy sample (see Section 2.3). We now discuss the steps of this procedure in turn.

4.1 Age distribution matching

In this section, we describe our two-phase algorithm for assigning both luminosity and colour to the galaxies located at the centre of (sub)haloes in an *N*-body simulation. We describe the luminosity assignment phase in Section 4.1.1, the colour assignment phase in Section 4.1.2 and provide a summary in Section 4.1.3.

4.1.1 Luminosity assignment

We assign *r*-band luminosity to Bolshoi (sub)haloes using the standard abundance matching technique. This technique has been widely used in the literature and so we do not discuss all of the technical details of this phase of our algorithm. Our implementation of the abundance matching luminosity assignment is identical to that described in detail in appendix A of H13, and thus we restrict ourselves to an overview here.

---

[1] See Skibba et al. (2011) for a discussion of the complicating factor that central galaxies are not always the brightest group members.

The quantity $V_{max} \equiv \text{Max} \left\{ \sqrt{GM(<r)/r} \right\}$, where $M(<r)$ is the mass enclosed within a distance r of the halo centre, defines the maximum circular velocity of a test particle orbiting in the halo's gravitational potential well. The abundance matching algorithm for luminosity assignment assumes a monotonic relationship between luminosity and V_{max}, such that the cumulative abundance of SDSS galaxies brighter than luminosity L, $n_g(>L)$, is equal to the cumulative abundance of haloes with circular velocity larger than V_{max}, $n_h(>V_{max})$. Following common practice, we abundance match on the property V_{peak}, the maximum value V_{max} ever attains over the entire merger history of the (sub)halo (see Reddick et al. 2013 for a thorough discussion of the use of V_{peak} in abundance matching).

Recent results (Wetzel & White 2010) indicate that haloes at the faint end of our mock galaxy sample ($M_r \lesssim -19$) may be pushing the resolution limit of the Bolshoi simulation. We do not attempt to correct for possible resolution effects, but note that such corrections (e.g. Moster et al. 2010) may be important for precision predictions of the small-scale 2PCF. A systematic investigation of these resolution effects is beyond the scope of this paper, and so we leave this as a task for future work.

Galaxy formation is a complex process, and so a single halo property such as V_{max} cannot uniquely specify the luminosity of the halo's galaxy. For example, the Tully–Fisher relation has intrinsic scatter (Pizagno et al. 2007), and so abundance matching cannot reproduce the statistics of the observed galaxy distribution without scatter (see Behroozi, Conroy & Wechsler 2010 for a comprehensive analysis of sources of scatter in abundance matching modelling). Our luminosity abundance matching technique results in ∼0.15 dex of scatter in luminosity at fixed V_{max} over the luminosity range of our sample, in accord with results from satellite kinematics (More et al. 2009) as well as other abundance matching studies (Reddick et al. 2013). Details of our implementation of scatter can be found in appendix A of H13.

4.1.2 Colour assignment

Once luminosities have been assigned to the mock galaxies, we proceed to the second phase of our algorithm. First, we divide the mock and SDSS galaxies into 10 evenly spaced bins of *r*-band magnitude. We use the observed colours of the SDSS Mr19 galaxies in each luminosity bin to define the probability distribution $P_{SDSS}(g - r|L)$. If there are N_L mock galaxies in luminosity bin L, we randomly draw N_L times from $P_{SDSS}(g - r|L)$ and rank-order the draws, reddest first. We assign these colours to the mock galaxies in luminosity bin L after first rank-ordering these N_L mock galaxies by the property z_{starve} (defined below), largest first.

By repeating the above procedure in each luminosity bin, the colour distribution of the mock galaxies is in exact agreement with the data, as can be seen in the top-left and bottom panels of Fig. 1. Note that this desirable property of our mock would be true *regardless* of the rank-ordering on z_{starve}. The effect of the rank-ordering is to introduce, at fixed luminosity, a correlation between galaxy colour and the epoch of starvation.

The property z_{starve} is based on three characteristic epochs of a halo's mass assembly history. These are as follows.

(1) z_{char}: the first epoch at which halo mass exceeds 10^{12} h^{-1} M$_\odot$. For haloes that never attain this mass, $z_{char} = 0$.

(2) z_{acc}: for subhaloes, z_{acc} is the epoch after which the object always remains a subhalo. For host haloes, $z_{acc} = 0$.

(3) z_{form}: using the methods of Wechsler et al. (2002), we identify the redshift at which the halo transitions from the fast- to
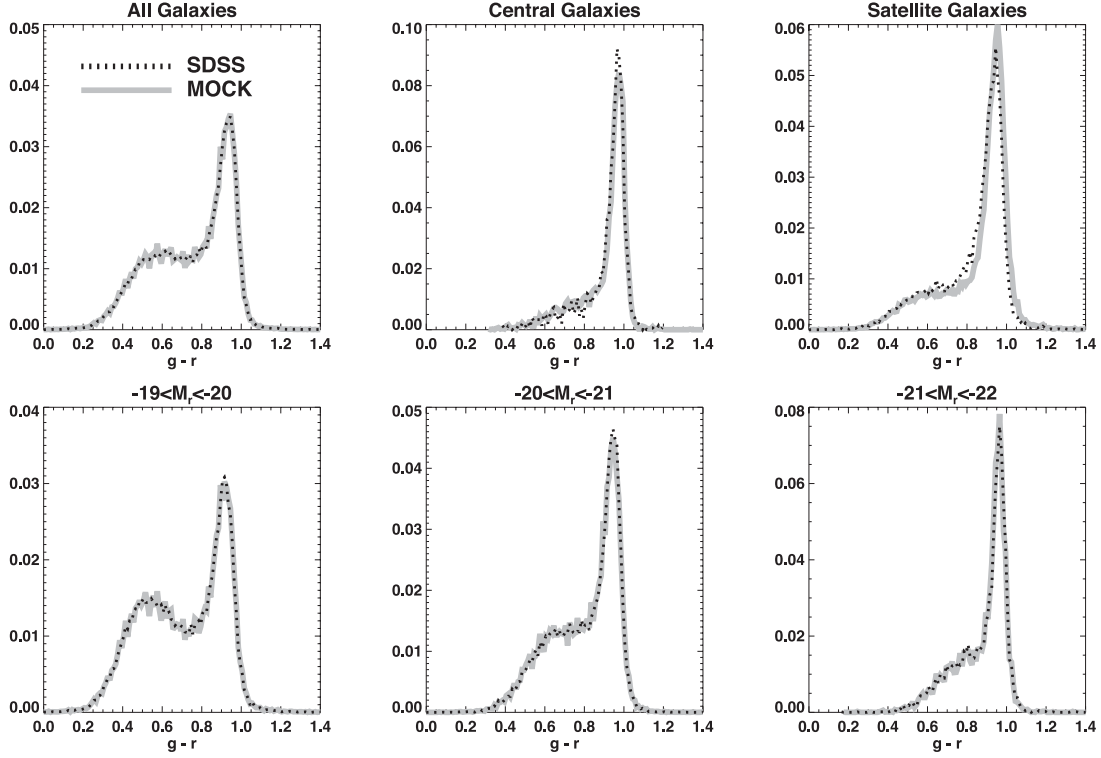
Figure 1. Galaxy colour PDFs from our mock catalogue as compared to those measured in the Mr19 SDSS galaxy catalogue. By construction, the $(g - r)$ probability distribution of our mock galaxies is in exact agreement with the data (black dotted histograms) at all luminosities. This is evident in the top-left panel and in all three bottom panels. Additionally, we compare colour PDFs for central and satellite galaxies in the group catalogue in the top-centre and top-right panels, respectively. Since information about satellite/central designation does not inform the colours we assign to the galaxies, the centre and right-hand panels on the top row illustrate successful *predictions* of distribution age matching.

slow-accretion regime, as this is the epoch after which dark matter ceases to accrete on to the halo's central region.

From these three characteristic epochs, we define the redshift of starvation:

$$z_{\text{starve}} \equiv \text{Max} \{ z_{\text{acc}}, z_{\text{char}}, z_{\text{form}} \} . \tag{2}$$

We provide a full discussion of the physical motivation and interpretation of z_{starve} in Section 6.3. In Appendix A, we discuss the details of how each of the above properties is computed from Bolshoi merger trees.

4.1.3 Algorithm summary

For convenience, we conclude discussion of our mock-making algorithm by briefly summarizing the steps discussed above.

(i) From the Bolshoi halo catalogue and merger trees, we compute V_{peak} and z_{starve} for every (sub)halo in the simulation.

(ii) We assign *r*-band luminosities to the (sub)haloes using abundance matching with the halo property V_{peak}. Mock galaxies dimmer than $M_r > -19$ are cut from the sample.

(iii) The mock galaxies are binned by luminosity, and colours are assigned by randomly drawing from the colour Probability Distribution Function (PDF) $P_{\text{SDSS}}(g - r | L)$ defined by our SDSS Mr19 galaxy sample. The random draws from $P_{\text{SDSS}}(g - r | L)$ are rank-ordered, and the mock galaxies in each luminosity bin are rank-ordered by the halo property z_{starve}, such that in each luminosity bin mock galaxies where z_{starve} occurs earlier are assigned redder colours.

4.2 Group identification

Once colours and luminosities have been assigned to the halo catalogue, we place our mock galaxies into redshift space and run the same group-finding algorithm on the mock that was run on the SDSS Mr19 galaxy sample. In this way, our mock groups inherit the same purity and incompleteness systematics as our SDSS groups. Finally, we introduce fibre collisions into our mock sample according to the prescription adopted in H13; this method preserves the relationship between the local galaxy density and the likelihood that a fibre collision occurs in the SDSS Mr19 sample.

5 RESULTS

We now investigate how well our model can predict a variety of observed galaxy statistics. First, we illustrate the general relationship between $(g - r)$ and z_{starve}, and then demonstrate that our model naturally predicts the correct relative colours of central and satellite galaxies. We show the accuracy of our base Mr19 SDSS catalogue at reproducing the observed SDSS luminosity-binned 2PCF, explicitly showing that our mock catalogue naturally inherits the successes of abundance matching. We then compare the resulting colour-binned 2PCF predictions to SDSS measurements. Finally, we test the success of our model against a variety of galaxy group statistics measured from the Mr19 group catalogue.

5.1 Halo age and galaxy colour

In Fig. 2, we plot the mean z_{starve} value of (sub)haloes in our mock in bins of $(g - r)$ colour. The grey band shows the result for all
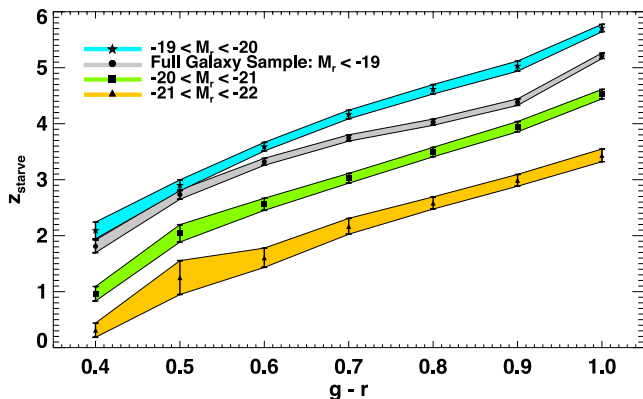
Figure 2. Implied relation between the $g - r$ colour and z_{starve}. Results for the $M_r < -19$ threshold sample appear in grey; luminosity-binned samples are coded according to the legend. The error bars are estimated by bootstrapping. The relation varies between luminosity bins due to the luminosity dependence on the colour PDF $P_{\text{SDSS}}(g - r|L)$. All results demonstrate the defining assumption of age distribution matching that galaxies with older stellar populations are typically found in older haloes.

mock galaxies in our sample, while the result in bins of luminosity appears in the coloured bands. The width of the bands corresponds to bootstrap estimates of the error on the mean. This figure visually demonstrates the fundamental assumption of age distribution matching: redder galaxies tend to live in older haloes. The luminosity dependence of the z_{starve}–colour relation simply reflects the colour–luminosity trend seen in the data, $P_{\text{SDSS}}(g - r|L)$.

The values of z_{starve} may seem strikingly large. Indeed, the reddest galaxies in our model typically have $z_{\text{starve}} \approx 5$–$6$. This highlights an important point: the property z_{starve} does not correspond to the redshift where star formation in the galaxy is quenched. Rather, z_{starve} signifies a special epoch in halo assembly that correlates strongly with the epoch when star formation in the galaxy becomes inefficient. In our colour assignment algorithm (described in Section 4.1.2), rank-ordering on z_{starve} only serves to govern the *relative* colours assigned to haloes; drawing the assigned colours from $P_{\text{SDSS}}(g - r|L)$ ensures that the absolute values of the assigned colours are correct. Thus, in the (likely) event that there is a substantial time delay between z_{starve} and z_{quench}, provided that this time delay does not strongly vary with halo mass, rank-ordering on z_{starve} should still suffice to assign the right colours to the right haloes. We return to this point in Section 6.

5.2 Central and satellite colours

As discussed in Section 4.1.2, our model correctly reproduces the colour distribution $P_{\text{SDSS}}(g - r|L)$ by construction. However, this by no means guarantees that our colour PDF will be correctly predicted when conditioned on some other property besides luminosity. In the top-centre and top-right panels of Fig. 1, we show the colour PDFs of our entire sample, conditioned on whether the galaxy is a central or satellite, respectively.

Our colour assignment only uses the property z_{starve} and $P_{\text{SDSS}}(g - r|L)$ to assign colours to the mock galaxies but does not distinguish between centrals and satellites. Thus, the good agreement seen between our model and the data in the top-centre and top-right panels of Fig. 1 demonstrates a successful *prediction* of age distribution matching. In our model, central and satellite galaxies have different colour distributions because host haloes and subhaloes have different mass assembly histories.

5.3 Luminosity- and colour-dependent clustering

We begin discussion of our clustering results by showing that our abundance matching-based luminosity assignment (described in Section 4.1.1) accurately predicts the observed luminosity-binned 2PCF as measured by Zehavi et al. (2011). To do this, we take our volume-limited Mr19 catalogue and subdivide into three luminosity bins: $-19 \le M_r \le -20$, $-20 \le M_r \le -21$ and $-21 \le M_r \le -22$. For each of these luminosity bins, we compute the real-space 2PCF. We then convert to projected space, $w_p(r_p)$, via

$$w_p(r_p) = 2 \int_0^{\pi_{\text{max}}} \xi\left(\sqrt{r_p^2 + \pi^2}\right) d\pi. \qquad (3)$$

For the upper limit in the integration, we use $\pi_{\text{max}} = 40\,h^{-1}\,\text{Mpc}$ to be consistent with the measurements made for our observational samples.

The black solid curves in the left column of Fig. 3 show the 2PCF as predicted by our mock catalogue. There is excellent agreement with the Zehavi et al. (2011) SDSS measurements (black data points) at each luminosity bin and at all projected separations from the linear to highly non-linear regimes.

From this success, we now turn to the right column of Fig. 3 showing the colour-dependent 2PCFs in distinct r-band luminosity bins as predicted by our age distribution matching formalism. The red filled and blue open circles in all panels represent the red and blue galaxy populations from Zehavi et al. (2011), respectively. The red and blue solid curves denote the model predictions. While there
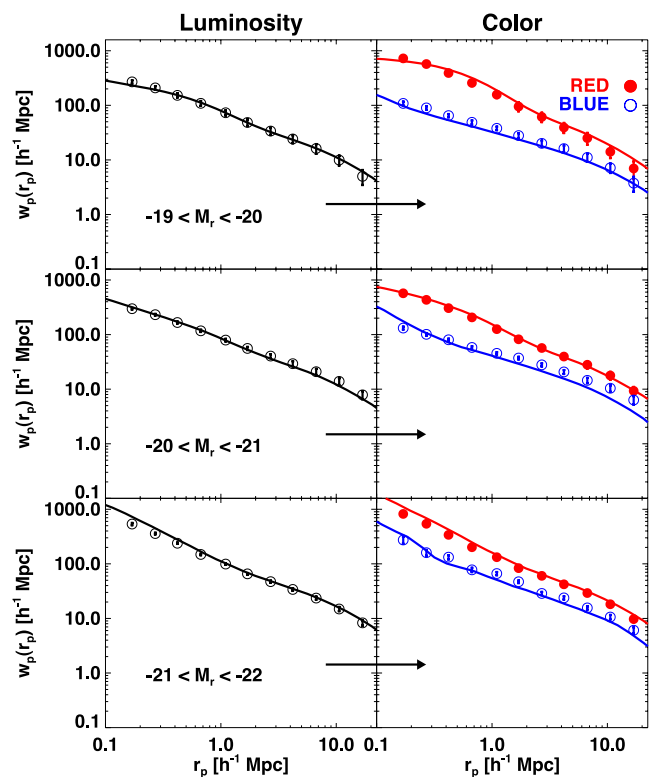

Figure 3. Luminosity- and colour-dependent clustering as predicted by our age distribution matching formalism. Left column: the luminosity-binned projected 2PCF predicted by our model (black solid curves) against the clustering exhibited by SDSS galaxies. Right column: in bins of luminosity, we plot the projected 2PCF of red (blue) mock galaxies with red (blue) solid curves. The red filled (blue open) points show the clustering of red (blue) SDSS galaxies. We use the results presented in Zehavi et al. (2011) for all SDSS measurements shown in this figure.
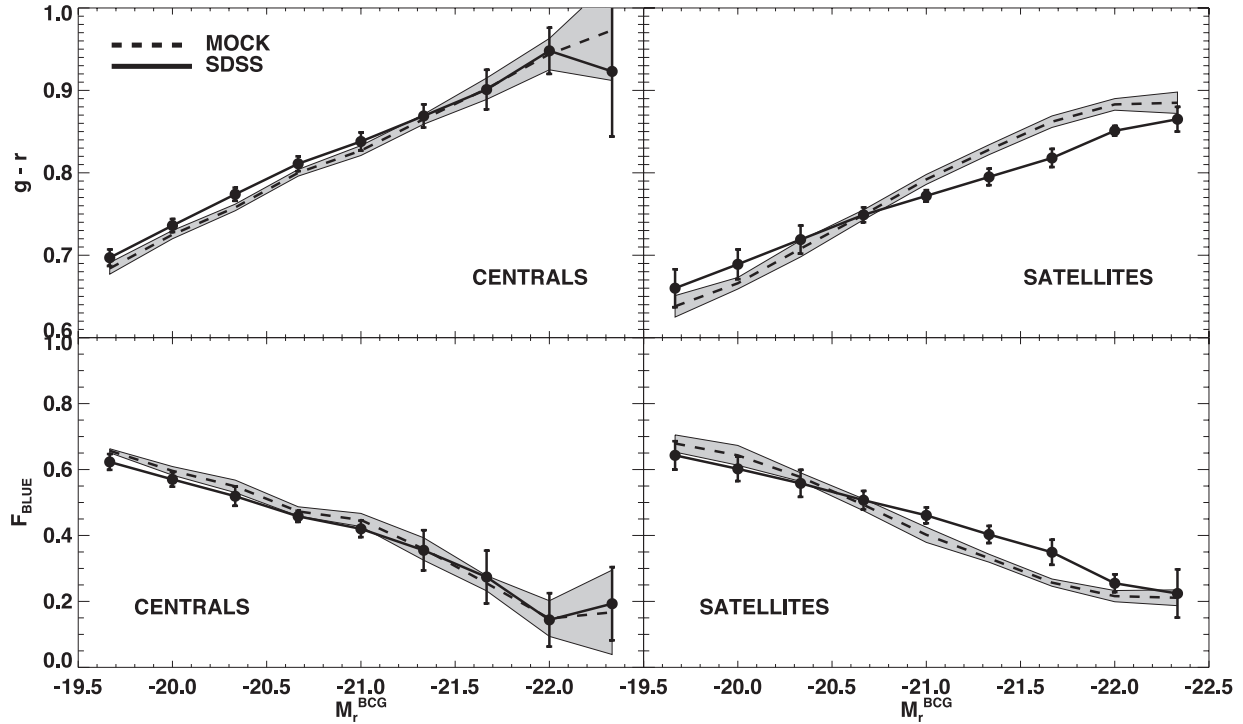
Figure 4. Top row: mean $(g - r)$ colour as a function of the absolute magnitude of the brightest central galaxy, M_r^{BCG}, for central galaxies (left-hand panel) and satellite galaxies (right-hand panel). SDSS data are shown as the black filled circles and the dashed curves represent the predictions from our mock catalogue. The solid grey bands denote the errors computed as the dispersion over 1000 bootstrap realizations of the galaxy sample (see Section 5.4 for details). Bottom row: same as the top row but now with the fraction of galaxies in groups of a given M_r^{BCG} which are blue, F_{BLUE}, as defined by the colour cut in equation (1).

is some discrepancy between the clustering of blue galaxies in the model and the data at large scales for the $-20 \leq M_r \leq -21$ bin, overall there is good agreement for each luminosity bin and at all r_p separation scales. This result is very encouraging since the colour-dependent 2PCF encodes rich information about the galaxy–halo connection (e.g. Cacciato et al. 2013; More et al. 2013). However, as shown in Masaki et al. (2013), the 2PCF alone is not enough to discriminate between competing models for the relationship between a halo and the colour of the galaxy it hosts (see also Neistein et al. (2011b)). We address this in the next section with our exploration of galaxy group statistics.

We emphasize that we have done no parameter fitting to achieve the agreement between the predicted and measured colour-binned clustering. Our algorithm for colour assignment has no explicit dependence on halo position; rather, the clustering signal in our mock emerges as a *prediction* of our theory. Since the primary aim of this paper is to introduce the age distribution matching formalism, we have opted for the simplest implementation of this framework rather than attempting to finely tune our model according to the measured clustering (see Section 6 for a discussion of more complex possible implementations). The success of our model's prediction for the colour-binned 2PCF is rather striking given the simplicity of the formulation we present here.

5.4 Galaxy group statistics

In this section, we test how well our model distinctly assigns colours to satellite and central galaxies in our mock catalogue compared to those colours in the SDSS galaxy group catalogue. In the *top row* of Fig. 4, we show the mean $(g - r)$ colour of group galaxies as a function of the brightness of the group's central galaxy, M_r^{BCG}.

From left to right, we show the results for central galaxies and satellites, respectively, in the mock and the SDSS group catalogue. The dashed line traces the mean $(g - r)$ colour of the mock galaxies in a given M_r^{BCG} bin. The *bottom row* is similar to the top, except that we plot $F_{\mathrm{BLUE}}(L)$, the fraction of galaxies that are designated 'blue' according to the colour cut defined by equation (1).

In all panels of Fig. 4, the solid grey region shows the errors on the mean. We use standard bootstrap resampling techniques to estimate the errors in all of our group-based statistics. We compute the errors as the dispersion over 1000 bootstrap realizations of the galaxy sample, where each realization is constructed by randomly selecting[2] N_g galaxies from the sample, where N_g is the total number of galaxies in the sample.

The successful, distinct predictions for the colours of central and satellite galaxies emerge naturally from the age distribution matching formalism. At fixed luminosity, the colours of our mock galaxies are drawn from the same colour PDF, $P_{\mathrm{SDSS}}(g - r|L)$, regardless of sub/host halo designation. Moreover, our colour assignment algorithm takes no explicit account of (sub)halo mass. Thus, the difference between the colours of satellites and centrals in our mock is purely a reflection of the different mass assembly history of host haloes and subhaloes. Excepting only some mild tension for the colour of satellite galaxies in a few of the brighter M_r^{BCG} bins, the agreement between mock and observed central and satellite galaxy colour is quite good. Again, we emphasize that we have not tuned any parameters in our model and we have chosen to

---

[2] We bootstrap resample *with replacement*, that is, we allow for the possibility of repeated draws of the same object.

keep our implementation of the age distribution matching technique as simple as possible (see Section 6 for an elaboration of this point).

6 DISCUSSION AND IMPLICATIONS

6.1 Comparison to abundance matching

Mock catalogues based on age distribution matching inherit all the successes of abundance matching because this is precisely the method we use in the luminosity assignment phase of our algorithm. The colour PDF $P_{\mathrm{SDSS}}(g - r | L)$ plays the same role in the colour assignment that the luminosity function $\Phi_{\mathrm{SDSS}}(L)$ plays in the luminosity assignment. Thus, just as $\Phi_{\mathrm{mock}}(L)$ from abundance matching agrees precisely with the observed luminosity function, so do our mock colour distributions agree precisely with the observed colour PDF.

There is an important conceptual difference between abundance matching and age distribution matching: the role that the property z_{starve} plays in the colour assignment differs markedly from the role played by V_{max} in the luminosity assignment. In abundance matching, luminosity assignment simply cannot proceed without appeal to the property V_{max} (or some other halo property such as M_{vir} for host haloes and M_{peak} for subhaloes), because this is an essential ingredient to the defining relation $n_{\mathrm{g}}(>L) = n_{\mathrm{h}}(>V_{\mathrm{max}})$. This is not the case with the colour assignment. While the dark matter halo property z_{starve} is in monotonic correspondence with the $(g - r)$ colour just as V_{max} is in monotonic correspondence with luminosity, this feature is not essential in order for the colour assignment algorithm to proceed. Indeed, the rank-ordering of the $(g - r)$ colour with z_{starve} only serves to correlate galaxy colour with halo mass assembly history (at fixed V_{max}), but a mock catalogue constructed without this correlation would nonetheless have a colour PDF that is in agreement with $P_{\mathrm{SDSS}}(g - r | L)$.

In light of this observation, one may wonder whether correlating galaxy colour with z_{starve} is necessary at all in the construction of a successful model. However, a mock catalogue constructed by drawing colours from the data without any rank-ordering predicts no difference between the clustering of red and blue samples (at fixed luminosity). This failure is easy to understand: in such a model, colours are assigned purely at random, and so there is no mechanism which produces any trend between colour and the cosmic density field.

6.2 Comparison to HOD models

In the standard implementation of the HOD formalism, the only halo property that governs galaxy occupation is the host halo mass, M_{host}. In particular, both Zehavi et al. (2011) and Skibba (2009) employ the HOD in their distinct approaches to modelling colour-dependent clustering, and in neither model is there any explicit use of halo age. In the Zehavi et al. (2011) model, the HOD parameters of red and blue galaxy samples are constrained by two-point clustering together with galaxy abundance. In the alternative HOD approach explored in Skibba (2009), the authors draw $(g - r)$ colours for mock centrals and satellites from explicitly different distributions. Both models faithfully reproduce the colour dependence of the 2PCF [see also Skibba et al. (2013) for tests of colour HOD models beyond binned 2PCF statistics].

The above HOD-based models assume that there are only two pieces of information that determine (in a statistical sense) the colour of a galaxy: (1) whether the galaxy is a central or a satellite and (2) the mass of the host halo in which the galaxy resides. As both

of these models enjoy great success in reproducing observations, it may seem surprising that our model also succeeds since the colour assignment phase of our algorithm appears to take no account of host halo mass. Moreover, in our model the only distinction between centrals and satellites is the use of z_{acc}, yet $z_{\mathrm{starve}} \neq z_{\mathrm{acc}}$ for the majority (>90 per cent) of subhaloes at all luminosities. Thus, our model also appears to assign colours to central and satellite galaxies in an identical fashion, at least in effect.

The resolution to this puzzle is illuminating. First, in CDM there is a strong correlation between halo mass and formation, with more massive haloes forming at later times. Thus, our model effectively *does* introduce rather strong trends between the $(g - r)$ colour and host halo mass due to the halo age–mass correlation. Thinking of M_r^{BCG} as a proxy for host halo mass, these trends can readily be seen in Fig. 4.

Secondly, at fixed mass, satellites and centrals have quite different assembly histories. This is true even when accounting for post-accretion stripping by comparing centrals to satellites at fixed mass at accretion, M_{acc}. In general, at fixed M_{acc}, satellites formed earlier than centrals, since satellites formed in denser regions of the initial density field and rapidly accreted their mass at earlier times. In age distribution matching, it therefore naturally emerges that satellites and centrals have different colours (Fig. 1) even though we assign colours to them from the same PDF.

We conclude our discussion of models based only on host halo mass by noting that any model for galaxy colour that depends strictly on M_{host} is unable even in principle to reproduce the phenomenon of so-called *Galactic conformity*: at fixed host mass, group systems with a blue central have a bluer satellite population than groups with a red central (Weinmann et al. 2006). This feature emerges naturally in the age distribution matching formalism, as we will demonstrate explicitly in a future companion paper.

6.3 Physical motivation for z_{starve}

The expectation that z_{starve} should be related to stellar mass assembly is physically well motivated. First, it is natural to expect that z_{form} represents an important epoch in the history of the formation of a galaxy. As shown in Wechsler et al. (2002), at z_{form} the dark matter halo transitions from the fast- to slow-accretion regime. After this epoch, the matter accreting on to the dark matter halo ceases to penetrate into the core of the halo where the galaxy resides,[3] depriving the galaxy of new material from the field with which it can form stars.

In reality, the halo may continue to form stars from its existing gas supply after this epoch, leading to a delay between z_{form} and the actual quenching of star formation. However, our model is based on *rank-ordering* of the property z_{starve}, and so the *relative* colour of galaxies in haloes with different formation times will be correct provided that this time delay is not a strong function of redshift or halo mass. As discussed in Section 6.5, we will explore more complex models with quenching time delays in a future companion paper. We simply note here that corrections due to mass- and redshift-dependent time delays cannot be too severe or our model would not successfully predict the 2PCF and group-based statistics.

Secondly, the existence of a characteristic halo mass M_{char} above which star formation is highly inefficient also has strong empirical support. The halo mass dependence of the stellar mass-to-halo mass ratio of galaxies shows a peak at $M_{\mathrm{char}} \sim 10^{12} \, h^{-1} \, \mathrm{M}_{\odot}$ that rapidly

---

[3] See their fig. 18 for a demonstration of this point.

falls off as halo mass increases (Behroozi, Wechsler & Conroy 2013d; Yang et al. 2012, 2013; Moster, Naab & White 2013; Watson & Conroy 2013). This characteristic mass was recently shown to remain essentially constant for most of cosmic history (Behroozi, Wechsler & Conroy 2013a). Additionally, $M_{char} \sim 10^{12}\,h^{-1}\,M_\odot$ is the halo mass at which pressure-supported shocks can heat infalling gas to the virial temperature (Dekel & Birnboim 2006). More importantly, this is also the mass when effects due to AGN feedback become significant, which can have a dramatic effect on star formation (Shankar et al. 2006; Teyssier et al. 2011; Martizzi, Teyssier & Moore 2012).

In our implementation of age distribution matching, we frame no hypothesis for the particular physical mechanism(s) that influence star formation inside massive haloes. Instead, we simply posit that there exists a characteristic halo mass M_{char} above which star formation becomes inefficient, and that this leads to a correlation between the galaxy colour and the characteristic epoch z_{char} that the halo attains at this mass.

We explored a range of masses for M_{char} and find that both the 2PCF and our group statistics allow us to discriminate between competing values for the characteristic mass. In models where $M_{char} \gtrsim 10^{12.5}\,h^{-1}\,M_\odot$, bright central galaxies are far too blue, and bright blue galaxies are strongly over-clustered on small scales ($\lesssim 1\,h^{-1}$ Mpc). Conversely, in models with $M_{char} \lesssim 10^{11.5}\,h^{-1}\,M_\odot$, central galaxies are too red and the 2PCF is in poor agreement with the data in all luminosity bins.

Finally, we turn to the property z_{acc}, the epoch a satellite accreted on to its host halo. Numerous studies have shown that the dependence of galaxy colour and/or star formation rate (SFR) has little to no residual environmental dependence once small-scale environment ($\lesssim 1\,h^{-1}$ Mpc) has been accounted for (e.g. Hogg et al. 2004; Blanton et al. 2005). This small-scale environmental dependence has been shown to be driven primarily by correlations between the properties of the host halo and its satellite galaxies (Blanton & Berlind 2007; Tinker, Wetzel & Conroy 2011; Wetzel et al. 2012, 2013a). These results have given rise to a widely accepted picture in which satellite evolution is significantly influenced by processes that occur near or within the virial radius of the host halo.

There are a variety of physical mechanisms that may have a significant impact on satellite galaxy colour/SFR. These include, but are not limited to, mass-loss due to outside-in stripping (see Watson et al. 2012b, and references therein), 'strangulation' (Larson, Tinsley & Caldwell 1980; Diemand, Kuhlen & Madau 2007; Kawata & Mulchaey 2008), ram-pressure stripping (Gunn & Gott 1972; Abadi, Moore & Bower 1999; McCarthy et al. 2008), 'harassment' (Moore, Lake & Katz 1998) and intra-host mergers (Makino & Hut 1997; Wetzel, Cohn & White 2009a,b). Again, in our implementation of age distribution matching, we do not preferentially select any one of these physical mechanisms as the primary cause for satellite quenching. Instead, we make the following simple observation: the impact that each of the scenarios discussed above has on satellite colour/SFR will be greater for satellites that fell into their host halo potential at earlier times. Thus, it is natural to expect a correlation between satellite galaxy colour and z_{acc}. Accordingly, we identify z_{acc} as the third 'special' epoch in (sub)halo mass accretion history.

Our treatment of the effect satellite accretion has on colour/SFR is consistent with the finding in Wetzel et al. (2012) that there is no minimum host halo mass for post-accretion physics to significantly influence satellite evolution. Moreover, since we correlate colour with z_{acc} regardless of the host halo mass, the success of our implementation of age distribution matching provides new supporting evidence for conclusions drawn in van den Bosch et al. (2008), who

found that satellite-specific processes are equally efficient in host haloes of all masses.

With the above physical picture in mind, in age distribution matching we make the following simple assumption: the earlier any one of the above processes begins in a halo, the redder its galaxy will be today. Mathematically, this is formulated by introducing a monotonic correspondence between the $(g - r)$ colour and $z_{starve} \equiv \mathrm{Max}\{z_{form}, z_{char}, z_{acc}\}$, as described in Section 4.1.2.

In alternative models that neglect to account for z_{char} and instead define $z_{starve} \equiv \mathrm{Max}\{z_{form}, z_{acc}\}$, the clustering prediction for bright galaxies is in very poor agreement with the data. It is natural to expect that this failure would occur at the bright end: in such a model, group/cluster BCGs (brightest central galaxies) are all bright blue! This is because the haloes hosting these BCGs are very massive and therefore very late forming, and so without z_{char} these haloes have the smallest values of z_{starve} in the halo catalogue. Conversely, models that neglect to account z_{form} perform well at the bright end but fail to correctly predict the clustering of dimmer galaxies. Thus, both z_{char} and z_{form} are essential ingredients to the success of age distribution matching.

The same is not true for z_{acc}: models which do not account for subhalo accretion make quite similar predictions to those that do. This is because subhaloes are typically not very massive objects that form very early on, and so for the typical subhalo $z_{starve} \gg z_{acc}$. This result may seem surprising since so much of the literature on star formation focuses squarely on the influence of satellite-specific quenching. Thus, the success of our current model suggests that too much importance may have been placed on post-accretion physics (see also Watson & Conroy 2013 for an alternative manifestation of this point). However, since there is likely a long time delay between z_{starve} and the actual quenching of star formation, we anticipate that subhalo accretion will prove to be more important when we more thoroughly explore models based on $z_{quench} \equiv z_{starve} + \Delta t_{delay}$. We return to this point in Section 6.5.

6.4 Comparison to other phenomenological models

In a recent paper closely related to this work, Masaki et al. (2013) paint a binary 'red' or 'blue' colour on to (sub)haloes that have first been assigned luminosities based on abundance matching. The authors explored a variety of different halo properties to use in the colour assignment. They found that several different choices result in successful predictions for the 2PCF, motivating their use of galaxy–galaxy lensing measurements to discriminate between competing models. Since precision galaxy–galaxy lensing measurements that have been binned simultaneously on both the r-band luminosity and $(g - r)$ colour were not available, the authors instead used measurements binned on morphology and assumed a perfect correspondence between red (blue) galaxies and early (late)-type morphological classification.

While the assumption that all red (blue) galaxies have early (late)-type morphology is a reasonable first approximation (Sheldon et al. 2004), we do not rely on it in this work. Instead, we employed galaxy group-based statistics to provide our additional discriminating power. Galaxy group membership yields information about halo occupation that is independent from 2PCF information. Group membership is necessarily identified in redshift space, and so group-based statistics incorporate information based on the velocity field, whereas projected clustering does not. Additionally, group statistics can probe halo occupation at higher masses than it is possible to do with small-scale clustering (see Hearin et al. 2013a, fig. 2). Nonetheless, in a future companion paper to this one, we will make

new galaxy–galaxy lensing measurements of SDSS galaxies binned on the $(g - r)$ colour to provide further tests of the success of the model we present here.

The SAM approach to predicting galaxy properties such as colour and star formation differs markedly from the theory presented here. SAMs also use the hierarchical build-up of dark matter haloes as a foundation, but from here SAMs model the myriad baryonic processes that take place inside haloes with a set of analytic formulae together with a large number of finely tuned parameters (White & Frenk 1991; Cole et al. 1994; Kauffmann et al. 1999; Somerville & Primack 1999; De Lucia & Blaizot 2007; Guo et al. 2011). Such approaches have been quite successful at reproducing many statistics of the observed galaxy distribution, though the colour-binned 2PCF has been notoriously difficult to correctly predict in quantitative detail (see, e.g., figs 20 and 21 of Guo et al. 2011).

There have been several recent advances in the literature on predicting galaxy colour/SFR that dramatically simplify the set of assumptions and number of parameters required by most SAMs. Mutch et al. (2013) couple dark matter halo accretion histories from the Millennium Simulation (Springel et al. 2005) with two simple analytic functions designed to encode the physics of baryonic growth and star formation history inside haloes (see also Cattaneo et al. 2011 for a closely related formulation). Wang et al. (2007) also use halo accretion histories in the Millennium Simulation together with a model of stellar population synthesis to paint galaxies with spectral energy distributions and SFR on to dark matter haloes; after fitting for two quenching time parameters, their model enjoys similar success to ours at predicting two-point clustering statistics.

Peng et al. (2010, 2012) begin with a minimal set of phenomenological observations based on zCOSMOS data (Lilly et al. 2007) and apply a set of continuity equations to the evolution of the number of red and blue galaxies in order to connect galaxy properties to the dark matter haloes hosting them (see also Lilly et al. 2013). The 'separability condition' discussed in Peng et al. (2010) is reflected by our formulation of the property z_{starve}, in that z_{char}, z_{form} and z_{acc} are calculated purely independently for each halo.

The studies discussed above (that is, Peng et al. 2012; Lilly et al. 2013; Mutch et al. 2013) assume a small set of simple analytic formulae together with knowledge of structure formation in ΛCDM to accurately reproduce a wide variety of one-point statistics of the galaxy distribution. By contrast, age distribution matching uses one-point statistics as input[4] and correctly reproduces the two-point function (Fig. 3) as well as higher order statistics based on galaxy groups (Fig. 4). The key to the success of our theory is the identification of a special time z_{starve} in each halo's assembly history that is physically associated with the slowing and ultimate quenching of star formation in the galaxy at the centre of the halo.

6.5 Future work

The primary aim of this paper is to introduce a new theoretical framework for studying the galaxy–halo connection, and to show that this model is successful even with our relatively simple implementation. Accordingly, wherever possible we have opted to employ simplifying approximations that are widely used in the literature rather than tuning our exploration of halo mass accretion history to provide better fits to the observations. For example, as discussed in Appendix A we assume that the formation history of

dark matter haloes can be encompassed by a single parameter, z_{form}, and we appeal to the results in Wechsler et al. (2002) in our use of halo concentration as a proxy for z_{form}. In this section, we call direct attention to some of these simplifying assumptions, and discuss possible extensions of our model that we intend to explore in future work.

6.5.1 Consideration of time delays

As discussed at length in Wetzel et al. (2012), the robust bimodality of the specific star formation rate distribution at low redshift constitutes compelling evidence that post-accretion quenching of satellite star formation occurs rapidly after substantial time delay (∼2–4 Gyr). This motivates a modification to the formalism we introduce here, in which rather than z_{acc}, the $(g - r)$ colour of environment-quenched satellites is instead correlated with $z_{\text{quench}} \equiv z_{\text{acc}} + \Delta t_{\text{delay}}$. We have conducted a preliminary investigation of such models, finding that $\Delta t_{\text{delay}} = 2$ Gyr has little effect on either the 2PCF or the group statistics. However, the physical processes leading to a correlation between colour and z_{char} may also have a delayed onset, for example due to the length of time required for a pressure-supported shock to propagate to the virial radius. Since the physics underlying these two delays ($\Delta t_{\text{delay}}^{\text{char}}$ and $\Delta t_{\text{delay}}^{\text{acc}}$) is likely quite distinct, there is no reason to assume that they are equal. We postpone a more systematic investigation of this parameter space for the companion paper mentioned above.

6.5.2 Implementation of Scatter

In abundance matching, there is scatter between luminosity and V_{max}. In this work, we have assumed a purely monotonic relation between the $(g - r)$ colour and the property z_{starve}, but it is easy to imagine implementations of age distribution matching in which there is scatter in this relation as well. A simple way to accomplish this would be an adaptation of the scatter method detailed in appendix A of Hearin et al. (2013a). Briefly, one would simply introduce Gaussian noise in the randomly drawn $(g - r)$ colours and use the rankings of the noisy $(g - r)$ values to appropriately shuffle the z_{starve}-ranking of the (sub)haloes. Again, we have not pursued this complication because our present aim is to show that even the simplest formulation of age distribution matching is remarkably successful. However, we intend to explore such extensions in future investigations of this technique, particularly when colour-binned galaxy–galaxy lensing measurements become available.

6.5.3 Generalizations of the model

In this paper, we have only investigated our theory at low redshift. However, if the luminosity function and colour PDF are well measured at other epochs, then our model should be consistent across redshifts, and this is a crucial next step for testing the generality of our age distribution matching formalism. Furthermore, our formalism generalizes in the obvious way to mock galaxies with stellar mass and/or (S)SFR instead of luminosity. We intend to study mocks based on such properties in the future as these are fundamental galaxy properties and are more directly linked to colour (for example, due to dust attenuation). Such tests will be enabled by existing 2PCF measurements as a function of stellar mass split on colour at $z \sim 0$ (Li et al. 2006; Jiang, Hogg & Blanton 2012) and also $z \sim 1$ (Mostek et al. 2013).

---

[4] Specifically, the luminosity function $\Phi_{\text{SDSS}}(L)$ and the distribution of colours $P_{\text{SDSS}}(g - r|L)$.

Our model is also general in the sense that the first phase of the algorithm need not necessarily be abundance matching. For example, one could instead begin by using the CLF formalism to assign brightnesses to dark matter haloes. Alternatively, a stellar mass-to-halo mass map could be applied to construct a mock based on stellar masses. In either case, the colour assignment phase of the algorithm would proceed in an identical fashion to what is described in Section 4.1.2.

Finally, we note that in this work we have not explicitly illustrated the relationships between z_{starve}, halo mass and luminosity (stellar mass), nor have we presented the relative importance of z_{acc}, z_{form} and z_{char} as a function of halo mass and luminosity (stellar mass). Such an investigation is of direct physical interest and may yield important insights into galaxy formation and evolution. We intend to conduct a detailed study of these relationships in future work once we have thoroughly explored more complex implementations of the age distribution matching technique along with the inclusion of additional observational constraints.

7 SUMMARY

In this section, we summarize our primary results and conclusions.

(1) We introduce *age distribution matching*, a new theoretical framework for studying the relationship between halo mass accretion history and stellar mass assembly.

(2) By construction, over the entire luminosity range of our sample the $(g - r)$ colour PDF of mock galaxies based on age distribution matching is in exact agreement with $P_{SDSS}(g - r|L)$.

(3) Our model successfully predicts luminosity- and colour-binned two-point projected galaxy clustering at low redshift, as well as central and satellite galaxy colour as a function of host halo environment (M_r^{BCG}). We emphasize that these are *predictions* of the theory, and that we have not relied upon any fitting or fine-tuning of any model parameters.

(4) The success of our model implies that the assembly histories of galaxies and haloes are indeed correlated.

(5) We make publicly available

(a) Our $M_r < -19$ mock galaxy catalogue.
(b) Full mass accretion histories of the main progenitors of all $z = 0$ Bolshoi haloes.

REFERENCES

Abadi M. G., Moore B., Bower R. G., 1999, MNRAS, 308, 947
Abazajian K. N. et al., 2009, ApJS, 182, 543
Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, ApJ, 600, 681
Balogh M. L., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1999, ApJ, 527, 54
Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Behroozi P. S., Wechsler R. H., Conroy C., 2013a, ApJ, 762, L31
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013b, ApJ, 762, 109
Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013c, ApJ, 763, 18
Behroozi P. S., Wechsler R. H., Conroy C., 2013d, ApJ, 770, 57
Bell E. F. et al., 2004, ApJ, 608, 752
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Berlind A. A. et al., 2006, ApJS, 167, 1
Blanton M. R., Berlind A. A., 2007, ApJ, 664, 791
Blanton M. R. et al., 2003, ApJ, 594, 186
Blanton M. R., Eisenstein D., Hogg D. W., Schlegel D. J., Brinkmann J., 2005, ApJ, 629, 143
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nat, 311, 517
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Cacciato M., van den Bosch F. C., More S., Mo H., Yang X., 2013, MNRAS, 430, 767
Carollo C. M. et al., 2012, preprint (arXiv:1206.5807)
Cattaneo A., Mamon G. A., Warnick K., Knebe A., 2011, A&A, 533, A5
Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781
Conroy C., Wechsler R. H., 2009, ApJ, 696, 620
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Cooper M. C. et al., 2006, MNRAS, 370, 198
Cooper M. C. et al., 2012, MNRAS, 419, 3018
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859
Gerke B. F., Wechsler R. H., Behroozi P. S., Cooper M. C., Yan R., Coil A. L., 2013, ApJS, 208, 1
Gottloeber S., Klypin A., 2008, preprint (arXiv:0803.4343)
Gunn J. E., Gott J. R., III, 1972, ApJ, 176, 1
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Guo Q. et al., 2011, MNRAS, 413, 101
Guo H., Zehavi I., Zheng Z., 2012, ApJ, 756, 127
Hearin A. P., Zentner A. R., Newman J. A., Berlind A. A., 2013a, MNRAS, 430, 1238
Hearin A. P., Zentner A. R., Berlind A. A., Newman J. A., 2013b, MNRAS, 433, 659 (H13)
Hogg D. W. et al., 2004, ApJ, 601, L29
Jiang T., Hogg D. W., Blanton M. R., 2012, ApJ, 759, 140
Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188
Kawata D., Mulchaey J. S., 2008, ApJ, 672, L103
Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Krause E., Hirata C. M., Martin C., Neill J. D., Wyder T. K., 2013, MNRAS, 428, 2548
Kravtsov A. V., 2013, ApJ, 764, L31
Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73
Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692

Leauthaud A. et al., 2012, ApJ, 744, 159

Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z., 2006, MNRAS, 368, 21

Lilly S. J. et al., 2007, ApJS, 172, 70

Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119

Makino J., Hut P., 1997, ApJ, 481, 83

Martizzi D., Teyssier R., Moore B., 2012, MNRAS, 420, 2859

Masaki S., Lin Y.-T., Yoshida N., 2013, preprint (arXiv:1301.1217)

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

Moore B., Lake G., Katz N., 1998, ApJ, 495, 139

More S., van den Bosch F. C., Cacciato M., Mo H. J., Yang X., Li R., 2009, MNRAS, 392, 801

More S., van den Bosch F. C., Cacciato M., More A., Mo H., Yang X., 2013, MNRAS, 430, 747

Mostek N., Coil A. L., Cooper M. C., Davis M., Newman J. A., Weiner B., 2013, ApJ, 767, 89

Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903

Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121

Mutch S. J., Croton D. J., Poole G. B., 2013, preprint (arXiv:1304.2774)

Neistein E., Weinmann S. M., Li C., Boylan-Kolchin M., 2011a, MNRAS, 414, 1405

Neistein E., Li C., Khochfar S., Weinmann S. M., Shankar F., Boylan-Kolchin M., 2011b, MNRAS, 416, 1486

Norberg P. et al., 2002, MNRAS, 332, 827

Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144

Peng Y.-j. et al., 2010, ApJ, 721, 193

Peng Y.-j., Lilly S. J., Renzini A., Carollo M., 2012, ApJ, 757, 4

Pizagno J. et al., 2007, AJ, 134, 945

Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, 771, 30

Riebe K. et al., 2011, preprint (arXiv:1109.0003)

Rodríguez-Puebla A., Drory N., Avila-Reese V., 2012, ApJ, 756, 2

Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, ApJ, 546, 20

Seljak U., 2000, MNRAS, 318, 203

Shankar F., Lapi A., Salucci P., De Zotti G., Danese L., 2006, ApJ, 643, 14

Sheldon E. S. et al., 2004, AJ, 127, 2544

Simha V., Weinberg D., Dave R., Fardal M., Katz N., Oppenheimer B. D., 2012, MNRAS, 423, 3458

Skibba R. A., 2009, MNRAS, 392, 1467

Skibba R. A., Sheth R. K., 2009, MNRAS, 392, 1080

Skibba R. A., van den Bosch F. C., Yang X., More S., Mo H., Fontanot F., 2011, MNRAS, 410, 417

Skibba R. A., Sheth R. K., Croton D. J., Muldrew S. I., Abbas U., Pearce F. R., Shattow G. M., 2013, MNRAS, 429, 458

Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087

Springel V. et al., 2005, Nat, 435, 629

Tasitsiomi A., Kravtsov A. V., Wechsler R. H., Primack J. R., 2004, ApJ, 614, 533

Teyssier R., Moore B., Martizzi D., Dubois Y., Mayer L., 2011, MNRAS, 414, 195

Tinker J. L., Wetzel A. R., 2010, ApJ, 719, 88

Tinker J., Wetzel A., Conroy C., 2011, preprint (arXiv:1107.5046)

Vale A., Ostriker J. P., 2004, MNRAS, 353, 189

Vale A., Ostriker J. P., 2006, MNRAS, 371, 1173

van den Bosch F. C. et al., 2007, MNRAS, 376, 841

van den Bosch F. C., Aquino D., Yang X., Mo H. J., Pasquali A., McIntosh D. H., Weinmann S. M., Kang X., 2008, MNRAS, 387, 79

Wang L., Li C., Kauffmann G., De Lucia G., 2007, MNRAS, 377, 1419

Watson D. F., Conroy C., 2013, ApJ, 772, 139

Watson D. F., Berlind A. A., McBride C. K., Hogg D. W., Jiang T., 2012a, ApJ, 749, 83

Watson D. F., Berlind A. A., Zentner A. R., 2012b, ApJ, 754, 90

Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52

Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, MNRAS, 366, 2

Weinmann S. M., Kauffmann G., van den Bosch F. C., Pasquali A., McIntosh D. H., Mo H., Yang X., Guo Y., 2009, MNRAS, 394, 1213

Wetzel A. R., White M., 2010, MNRAS, 403, 1072

Wetzel A. R., Cohn J. D., White M., 2009a, MNRAS, 394, 2182

Wetzel A. R., Cohn J. D., White M., 2009b, MNRAS, 395, 1376

Wetzel A. R., Tinker J. L., Conroy C., 2012, MNRAS, 424, 232

Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013a, MNRAS, 432, 336

Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013b, preprint (arXiv:1303.7231)

Whitaker K. E. et al., 2011, ApJ, 735, 86

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wyder T. K. et al., 2007, ApJS, 173, 293

Yang X., Mo H. J., van den Bosch F. C., 2003, MNRAS, 339, 1057

Yang X., Mo H. J., van den Bosch F. C., Zhang Y., Han J., 2012, ApJ, 752, 41

Yang X., Mo H. J., van den Bosch F. C., Bonaca A., Li S., Lu Y., Lu Y., Lu Z., 2013, ApJ, 770, 115

York D. G. et al., 2000, AJ, 120, 1579

Zehavi I. et al., 2002, ApJ, 571, 172

Zehavi I. et al., 2005, ApJ, 630, 1

Zehavi I. et al., 2011, ApJ, 736, 59

Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760

APPENDIX A

The central tenet of age distribution matching is that galaxy colour is determined by halo mass accretion history, at least in a statistical sense. In this appendix, we discuss the details of how we calculate z_{acc}, z_{form} and z_{char}, the key quantities used in the colour assignment phase of our algorithm.

z_{acc}: in accordance with the convention adopted in Behroozi et al. (2013c), we use the snapshot after which the subhalo always remains a subhalo. We also explored the definition of accretion advocated in Wetzel et al. (2013b), where subhaloes are said to have accreted the first time they have been identified as subhaloes for two consecutive time-steps, and find that this choice has little impact on our results.

z_{form}: we use the approximation introduced in Wechsler et al. (2002), who find that mass accretion history is well fitted by an exponential, $M_{vir}(z) = M_0 e^{-\alpha z}$. Wechsler et al. (2002) defined the scale factor of halo formation a_{form} to be the epoch when the logarithmic slope S of $M_{vir}(z)$ drops below $S = 2$. While there is some amount of arbitrariness in this particular chosen value of S, the transition that occurs at or near this time has very real physical consequences: the central density of the halo ceases to increase (see their fig. 18). This physical change naturally gives rise to a correlation between halo formation time and NFW concentration c_{vir} (defined as the ratio R_{vir}/r_s).

One of the central results of Wechsler et al. (2002) is that a_{form} is well approximated by the equation:

$$a_{form} \equiv 4.1 a_0 / c_{vir},$$

where $z_{form} = 1/a_{form} - 1$ and a_0 is the scale factor at the time of 'observation'. For host haloes we take the time of observation to be $a_0 = 1$ and use the halo's concentration c_{vir} at this snapshot. Tidal disruption and mass stripping of subhaloes after accretion make $a_0 = 1$ estimates of R_{vir} noisy and unreliable. Fortunately, Wechsler et al. (2002) showed that the scaling of halo formation with c_{vir} remains the same over a broad range of a_0. So for subhaloes, we define c_{vir} to be the NFW concentration of the subhalo at $a_0 = a_{acc}$, the time of accretion (as defined above). We note that for ~10 per

cent of subhaloes, this definition produces an unphysically long tail at high redshift in the z_{form} distribution. We attribute this to large, transient boosts to the subhalo concentration at a_{acc} due to a tidal event. However, we find that our results are insensitive to this behaviour: when we collapse all objects in this tail to $z_{\mathrm{form}} = 3$, we get virtually identical results.

z_{char}: we use the first time a (sub)halo attains $10^{12}\,h^{-1}\,\mathrm{M}_{\odot}$. As noted in the main body of the paper, we explored a range of values of M_{char} and found that $10^{12}\,h^{-1}\,\mathrm{M}_{\odot}$ is the most successful.

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