Testing subhalo abundance matching in cosmological smoothed particle hydrodynamics simulations

Vimal Simha,1,2⋆ David H. Weinberg,1⋆ Romeel Davel,3⋆ Mark Fardal,4⋆ Neal Katz,4⋆ and Benjamin D. Oppenheimer5⋆

1Astronomy Department and Center for Cosmology and AstroParticle Physics, Ohio State University, Columbus, OH 43210, USA
2Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE
3Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
4Astronomy Department, University of Massachusetts at Amherst, MA 01003, USA
5Leiden Observatory, Leiden University, 2300 RA Leiden, the Netherlands

Accepted 2012 April 20. Received 2012 April 19; in original form 2012 January 4

ABSTRACT
Subhalo abundance matching (also known as SHAM) is a technique for populating simulated dark matter distributions with galaxies, assuming a monotonic relation between a galaxy’s stellar mass or luminosity and the mass of its parent dark matter halo or subhalo. We examine the accuracy of SHAM in two cosmological smoothed particle hydrodynamics (SPH) simulations, one of which includes momentum-driven galactic winds. The SPH simulations indeed show a nearly monotonic relation between stellar mass and halo mass provided that, for satellite galaxies, we use the mass of the subhalo at the epoch $z_{\text{sat}}$ when it became a satellite. In each simulation, the median relation for central and satellite galaxies is nearly identical, though a somewhat larger fraction of satellites is outliers because of stellar mass loss. SHAM-assigned masses (at $z = 0–2$), luminosities ($R$-band at $z = 0$) or star formation rates (at $z = 2$) have a 68 per cent scatter of 0.09–0.15 dex relative to the true simulation values. When we apply SHAM to the subhalo population of a collisionless $N$-body simulation with the same initial conditions as the SPH runs, we find generally good agreement for the halo occupation distributions and halo radial profiles of galaxy samples defined by thresholds in stellar mass. However, because a small fraction of SPH galaxies suffer severe stellar mass loss after becoming satellites, SHAM slightly overpopulates high-mass haloes; this effect is more significant for the wind simulation, which produces galaxies that are less massive and more fragile. SHAM recovers the two-point correlation function of the SPH galaxies in the no-wind simulation to better than 10 per cent at scales $0.1 < r < 10 h^{-1} \text{Mpc}$. For the wind simulation, agreement is better than 15 per cent at $r > 2 h^{-1} \text{Mpc}$, but overpopulation of massive haloes increases the correlation function by a factor of $\sim 2.5$ on small scales. The discrepancy in the wind simulation is greatly reduced if we raise the stellar mass threshold from $6 \times 10^{9}$ to $3 \times 10^{10} \text{M}_\odot$; in this case SHAM overpredicts the SPH galaxy correlation function by $\sim 20$ per cent at $r < 1 h^{-1} \text{Mpc}$ but agrees well with SPH clustering at larger scales.

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

1 INTRODUCTION
In the standard theoretical description of galaxy formation, galaxies form by the dissipation of the baryonic component within collisionless dark matter (DM) haloes (e.g. White & Rees 1978; Fall & Efstathiou 1980). When a DM halo enters the virial radius of a more massive halo, it is subjected to tidal stripping and potentially to disruption. Nonetheless, high-resolution $N$-body simulations have indicated that massive haloes retain a substantial amount of substructure (Klypin et al. 1999; Moore et al. 1999; Springel, Yoshida & White 2001), consisting of bound DM clumps orbiting within the potential of their host halo. Evidently, such subhaloes were themselves independent, self-contained haloes in the past, before merging with a more massive halo. If sufficiently massive, these

*E-mail: vsimha@astronomy ohio-state.edu (VS); dhw@astronomy ohio-state.edu (DHW); rad@as.arizona.edu (RD); fardal@kaka.astro.umass.edu (MF); nsk@kaka.astro.umass.edu (NK); oppenheimer@strw.leidenuniv.nl (BDO)

© 2012 The Authors
Monthly Notices of the Royal Astronomical Society © 2012 RAS

Downloaded from https://academic.oup.com/mnras/article-abstract/423/4/3458/1747664 on 29 July 2018
subhaloes were sites of baryon dissipation and star formation in the past. There are many indications from studies of the statistical properties of how galaxies and substructures populate haloes that galaxies in groups and clusters are in fact the observational counterparts of subhaloes. For example, Colín et al. (1999) and Kravtsov et al. (2004) show that the correlation functions of substructures in high-resolution N-body simulations are in good agreement with the observed correlation functions of galaxies. On the theory side, Kravtsov et al. (2004) find that the distribution of subhaloes in high-resolution N-body simulations is similar to that of smoothed particle hydrodynamics (SPH) galaxies in Berlind et al. (2003) and Zheng et al. (2005).

Subhalo abundance matching (SHAM) is a technique for assigning observable galaxy properties to a halo/subhalo population in an N-body simulation. It is based on assuming a monotonic relationship between observable properties of galaxies and dynamical properties of DM substructures. As subhaloes that fall into the virial radius of more massive haloes are subjected to stripping and tidal disruption, several authors (e.g. Conroy, Wechsler & Kravtsov 2006; Vale & Ostriker 2006; Moster et al. 2010) contend that the properties of satellite galaxies should be better correlated with the properties of subhaloes at the time of their accretion onto a more massive halo rather than their present day properties. By using the SHAM technique, with this accretion-epoch matching, Conroy et al. (2006) match the observed luminosity dependence of galaxy clustering at a wide range of epochs, ranging from $z = 0$ to $\sim 5$. Assuming a monotonic relationship between galaxy mass and halo mass and using the same SHAM technique, Guo et al. (2010) reproduce the observationally inferred relation between stellar mass and halo virial mass. Trujillo-Gomez et al. (2011) use the abundance matching technique to match the observed relations between stellar mass and circular velocity and luminosity and circular velocity, and to match the estimated galaxy velocity function.

In this paper, we investigate the effectiveness of SHAM in cosmological SPH simulations, where we know exactly the relation between the properties of galaxies and the masses of their parent haloes and subhaloes. We investigate the degree to which there is a direct correspondence between the properties of DM substructures in a dissipationless numerical simulation of a cosmological volume and condensed baryons in numerical simulations (with the same initial conditions) that include a dissipative component. We extend the similar study of Weinberg et al. (2008) in several ways. First, our simulation volume is more than 10 times larger. Secondly, one of our dissipative simulations includes ejective feedback in the form of momentum-driven winds, which curtail star formation and produce a stellar mass function that is in better agreement with the observations. Lastly and perhaps most importantly, we relate galaxy properties to the subhalo mass at the epoch of accretion rather than at the present time (see comparison in Section 4). We make a direct assessment of the effectiveness of SHAM as a method for assigning stellar mass, luminosity or star formation rate (SFR) to subhaloes and investigate the sources of its breakdown.

Our investigation offers insight into the physical mechanisms that shape galaxy masses and luminosities in these simulations, and it also has practical import. If the subhalo population in the purely gravitational simulation does indeed trace the observable properties of the galaxy population, it enables us to make observable predictions of quantities like the luminosity dependence of galaxy clustering based on computationally less expensive N-body simulations instead of hydrodynamic simulations. SHAM also offers a relatively inexpensive tool for creating artificial galaxy catalogues to support statistical analyses of large-scale structure data sets.

In Section 2, we describe our simulation, our methods for identifying haloes, subhaloes and galaxies and the SHAM scheme. In Section 3, we investigate the relationship between the stellar masses, luminosities and SFRs of SPH galaxies and DM substructures in our SPH simulations to test some of the underlying assumptions of the SHAM technique. In Section 4, we investigate whether the SPH galaxy population can be recovered from the subhalo population in our matched N-body simulation; note that it is this comparison rather than the investigations in Section 3 that tests SHAM as it has been traditionally implemented. Finally, in Section 5, we summarize our results and discuss their implications.

2 METHODS

2.1 Simulations

Our simulations are performed using the GADGET-2 code (Springel 2005) as modified by Oppenheimer & Davé (2008). Gravitational forces are calculated using a combination of the particle-mesh algorithm (Hockney & Eastwood 1981) for large distances and the hierarchical tree algorithm (Barnes & Hut 1986; Hernquist 1987) for short distances. The SPH algorithm is entropy and energy conserving and is based on Springel & Hernquist (2002). The details of the treatment of radiative cooling can be found in Katz, Weinberg & Hernquist (1996). Gas can dissipate energy via Compton cooling and radiative cooling, computed assuming a primordial gas composition and a background ultraviolet (UV) flux based on Haardt & Madau (2001). The details of the treatment of star formation can be found in Springel & Hernquist (2003). Briefly, each gas particle satisfying a temperature and density criterion is assigned a SFR, but the conversion of gaseous material to stellar material proceeds stochastically. The parameters for the star formation model are selected so as to match the $z = 0$ relation between SFR and gas density (Schmidt 1959; Kennicutt 1998).

We adopt a $\Lambda$ cold dark matter ($\Lambda$CDM) cosmology (inflationary CDM with a cosmological constant) with $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$, $h = H_0/100$ km s$^{-1}$ Mpc$^{-1} = 0.7$, $\Omega_b = 0.044$, spectral index $n_s = 0.95$ and the amplitude of the mass fluctuations scaled to $\sigma_8 = 0.8$. These values are reasonably close to current estimates from the cosmic microwave background (Larson et al. 2004) and large-scale structure (Reid et al. 2010). We do not expect minor changes in the values of the cosmological parameters to affect our conclusions.

We follow the evolution of 288$^3$ DM particles and 288$^3$ gas particles, i.e. just under 50 million particles in total, in a comoving box that is 50 $h^{-1}$ Mpc on each side, from $z = 129$ to 0. The DM particle mass is $4.3 \times 10^8 M_\odot$, and the SPH particle mass is $9.1 \times 10^7 M_\odot$. The gravitational force softening is a comoving $5 h^{-1}$ kpc cubic spline, which is roughly equivalent to a Plummer force softening of 3.5 $h^{-1}$ kpc.

One of our simulations, SPH winds (SPHw) incorporates kinetic feedback through momentum-driven winds as implemented by Oppenheimer & Davé (2006) and Oppenheimer & Davé (2008) where the details of the implementation can be found. Briefly, wind velocity is proportional to the velocity dispersion of the galactic halo, and the ratio of the gas ejection rate to the SFR is inversely proportional to the velocity dispersion of the galactic halo.

We also carry out a simulation with the same cosmological and numerical parameters as the SPHw simulation, but without momentum-driven winds, SPH no winds (SPHnw). Although no energy is kinetically imparted to SPH particles, the SPHnw simulation includes thermal feedback from supernovae (see Springel & Hernquist 2003). However, since the surrounding gas is dense, the
energy is radiated away before it can drive outflows or significantly suppress star formation. Our SPHw and SPHnw simulations are analogous to the nw and vzw simulations of Oppenheimer et al. (2010), who investigate the growth of galaxies by accretion and wind recycling and compare predicted mass functions to observations. Our simulations are also used by Zu et al. (2011), who investigate intergalactic dust extinction.

In addition to the two SPH simulations, we carry out a non-dissipative, purely gravitational, N-body simulation with identical cosmological and numerical parameters and the same initial positions and velocities of particles as the SPH simulations, except that the DM particle mass is higher by a factor of $\Omega_m/(\Omega_m - \Omega_b)$ to compensate for the gravitational effects of not including baryons. We refer to this N-body simulation as the DM-only simulation.

### 2.2 Identification of groups and substructures

We identify DM haloes using a friends-of-friends (FOF) algorithm (Davis et al. 1985). The algorithm selects groups of particles in which each particle has at least one neighbour within a linking length. We assign a mass to the halo using a spherical overdensity (SO) criterion, with the threshold density set to the virial overdensity in spherical collapse (Kitayama & Suto 1996). At $z = 0$, the mean interior overdensity with respect to the critical density of the Universe is 94. We set the centre of the group at the most bound FOF particle and go out in radius until the mean density enclosed is equal to the virial density. While FOF occasionally links multiple distinct concentrations together, most FOF haloes contain a distinct mass concentration (see figs 1 and 2 of Simha et al. 2009).

To identify substructures within haloes, we use the ADAPTAHOP (Aubert, Pichon & Colombi 2004) code. The details of the algorithm can be found in Aubert et al. (2004), so we provide only a brief summary of it here. We calculate densities around each DM particle using an SPH-like kernel estimator, with a cubic spline kernel containing 32 neighbours. We then partition the ensemble of particles into ‘peak patches’ where a peak patch is a set of particles with the same local density maximum. The connectivity between peak patches is dictated by the saddle points in the density field. We identify subsets of particles in each peak patch with SPH density larger than the density of the highest saddle point connecting it to a neighbouring peak patch. To select statistically significant substructures as opposed to random fluctuations in the density field, we require that the mean density of a substructure be 4$\sigma$ above the SPH density of the saddle point. The performance of ADAPTAHOP on a simulation similar to that here is illustrated in Weinberg et al. (2008, see their fig. 1).

Hydrodynamic cosmological simulations that incorporate cooling and star formation produce dense groups of baryons with sizes and masses comparable to the luminous regions of observed galaxies (Katz 1992; Evrard, Summers & Davis 1994). We identify galaxies using the Spline Kernel Interpolative DENMAX (SKID)\(^1\) algorithm (Gelb & Bertschinger 1994; Katz et al. 1996), which identifies gravitationally bound particles associated with a common density maximum. We refer to the groups of stars and cold gas thus identified as galaxies. The simulated galaxy population becomes substantially incomplete below our resolution threshold of $\sim 64$ SPH particles (Murali et al. 2002), which corresponds to a baryonic mass of $5.8 \times 10^9 M_\odot$. Although this threshold applies to the total baryonic mass (stars plus cold, dense gas) of galaxies, we adopt it as our threshold for stellar mass and ignore galaxies with lower stellar mass.

Fig. 1 compares the galaxy stellar mass function in the SPHw and SPHnw simulations at $z = 0$. In the SPHnw simulation, there are 7952 galaxies above our resolution threshold, corresponding to a space density of 0.064 $h^3$ Mpc$^{-3}$. In the SPHw simulation there are only 2264 galaxies above our resolution threshold, corresponding to a space density of 0.018 $h^3$ Mpc$^{-3}$ because wind feedback pushes the stellar mass of many galaxies below the 64$ m_{SPH}$ threshold. The two mass functions gradually converge towards higher masses, joining at $M_S > 10^{11.8} M_\odot$, because wind feedback has less suppressing effect in larger systems since the amount of material ejected in this model scales inversely with circular velocity. Oppenheimer et al. (2010) discuss the comparison between the predicted stellar mass functions and observational estimates in some detail. Roughly speaking, the SPHw model reproduces observational estimates for $M_S < 10^{11} M_\odot$, but it predicts excessive galaxy masses (at a given space density) for $M_S > 10^{11.8} M_\odot$.

### 2.3 Subhalo abundance matching

SHAM is a technique for assigning galaxies to simulated DM haloes and subhaloes. The essential assumptions are that all galaxies reside in identifiable DM substructures and that luminosity or stellar mass of a galaxy is monotonically related to the potential well depth of its host halo or subhalo. Some implementations use the maximum of the circular velocity profile as the indicator of potential well depth, while others use halo or subhalo mass. The first clear formulations of SHAM as a systematic method appear in Conroy et al. (2006) and Vale & Ostriker (2006), but these build on a number of previous studies that either test the underpinnings of SHAM or implicitly assume SHAM-like galaxy assignment (e.g. Colín et al. 1999; Kravtsov et al. 2004; Nagai & Kravtsov 2005).

N-body simulations produce subhaloes that are located within the virial radius of SO haloes. The present mass of subhaloes is a product of mass build up during the period when the halo evolves in isolation and tidal mass loss after it enters the virial radius of a more massive halo (e.g. Kazantzidis et al. 2004; Kravtsov et al. 2004).

---

\(^1\) http://www-hpcc.astro.washington.edu/tools/skid.html
The stellar component, however, is at the bottom of the potential well and more tightly bound making it less likely to be affected by tidal forces. Therefore, several authors (e.g. Conroy et al. 2006; Vale & Ostriker 2006) argue that the properties of the stellar component should be more strongly correlated with the subhalo mass at the epoch of accretion rather than at \( z = 0 \).

Vale & Ostriker (2006) apply a global statistical correction to subhalo masses relative to halo masses (as do Weinberg et al. 2008), while Conroy et al. (2006) explicitly identify subhaloes at the epoch of accretion and use the maximum circular velocity at that epoch. Our formulation here is similar to that of Conroy et al. (2006), though we use mass rather than circular velocity. Specifically, we assume a monotonic relationship between stellar mass and halo mass and determine the form of this relation by solving the implicit equation

\[
   n_S(M_h) = n_M(M_{H}) \quad (1)
\]

where \( n_S \) and \( n_M \) are the number densities of galaxies and haloes, respectively, \( M_S \) is the galaxy stellar mass threshold and \( M_H \) is the halo mass threshold chosen so that the number density of haloes above it is equal to the number density of galaxies in the sample. The quantity \( M_H \) is defined as follows:

\[
   M_H = \begin{cases} 
   M_{h}(z = 0) & \text{for distinct haloes}, \\
   M_{h}(z = z_{\text{sat}}) & \text{for subhaloes},
   \end{cases}
\]

(2)

where \( z_{\text{sat}} \) is the epoch when a halo first enters the virial radius of a more massive halo. We also consider (in Section 3.2) variants of this procedure in which we substitute \( R \)-band luminosity or instantaneous SFR for stellar mass in equation (1).

In Section 3, we use our SPH simulations to test the degree to which galaxy stellar mass and luminosity are monotonic functions of halo/subhalo mass, as assumed in SHAM. For independent haloes, we set \( M_H \) equal to the \( z = 0 \) mass of SO haloes. We use the ADAPTAHOP code to identify subhaloes hosting satellite galaxies above the resolution threshold at \( z = 0 \) in our SPH simulations. We then use their particle membership to identify their progenitor SO haloes at \( z_{\text{sat}} \), whose mass we adopt as \( M_H \). We can identify a host subhalo for the vast majority of our galaxies at \( z = 0 \). However, for about 1 per cent of galaxies we cannot identify any associated substructure. In these cases, we use the galaxy’s baryonic particles to trace its high-redshift progenitors up to the epoch when it was the only galaxy in an SO halo and adopt the mass of this SO halo as \( M_H \). After matching, we examine the correlations of galaxy properties with \( M_H \) for both central and satellite systems, and we compare the stellar masses, luminosities and (at high redshift) SFRs that would be assigned by monotonic matching to the simulation values. Of course, SHAM is usually applied to collisionless \( N \)-body simulations, not SPH simulations, and the subhalo populations can differ even for the same initial conditions because of the dynamical effects of the dissipative baryons on the DM.

In Section 4, we populate haloes/subhaloes in our DM-only simulation using SHAM, following a procedure akin to the one usually used to match simulated haloes/subhaloes to observed galaxies. For independent haloes, we set \( M_H \) equal to the \( z = 0 \) mass of SO haloes. We identify subhaloes using the ADAPTAHOP code and use their particle membership to identify their progenitor SO haloes at \( z_{\text{sat}} \), whose mass we adopt as \( M_H \). We have checked that at \( z_{\text{sat}} \), the relation between \( v_{\text{max}} \) and \( M_H \) is close to monotonic, so we would get similar results from using \( v_{\text{max}} \) instead of \( M_H \). For comparison, we also implement a procedure that we refer to as SHAMz0, where we set \( M_H \) equal to the \( z = 0 \) mass for both independent haloes and subhaloes identified in our \( N \)-body simulation.

Knebe et al. (2011) find differences in the performance of various halo finding algorithms in identifying substructures and accurately recovering their properties, particularly for subhaloes containing fewer than 40 particles. Since we trace the \( z_{\text{sat}} \) progenitors of \( z = 0 \) substructures, our results are unlikely to be affected by random fluctuations in the density field of haloes that may be spuriously identified as substructures. However, if subhaloes hosting satellite galaxies that have merged with the central galaxy of the halo are identified as substructures, we would overestimate the halo occupation of massive haloes. Conversely, subhaloes that fall into more massive haloes and lose a substantial fraction of their mass due to tidal stripping may no longer be resolved in the simulation at \( z = 0 \) even though they might survive and host satellite galaxies in a sufficiently high-resolution calculation.

3 SHAM IN THE SPH SIMULATIONS

3.1 Stellar mass

The left-hand column of Fig. 2 shows galaxy stellar mass plotted against halo mass in the SPHnw simulation (top) and the SPHw simulation (bottom). For galaxies that are the central objects of their parent haloes (henceforth central galaxies), the halo mass, \( M_{H_1} \), on the horizontal axis, is the mass of the SO halo in which the galaxy is located at \( z = 0 \). For galaxies that are not the central objects of their respective SO haloes (henceforth satellite galaxies), \( M_{H_1} \) is the mass of the SO halo in which the galaxy is located at \( z_{\text{sat}} \), the last output epoch before its parent halo fell into the virial radius of a more massive halo: \( z_{\text{sat}} \) values range between \( z = 3 \) and 0.05. The black solid curve and red dotted curve show the median galaxy stellar mass in evenly spaced logarithmic bins of halo mass for all galaxies and satellite galaxies, respectively, while the red open circles and black open rectangles show individual galaxies that are within the top 5 per cent and bottom 5 per cent by stellar mass in each halo mass bin.

The key result of Fig. 2 is that the ratio of galaxy mass to \( M_H \) is similar for central galaxies and satellite galaxies. Between \( z_{\text{sat}} \) and \( z = 0 \), satellite galaxies have lower growth rates compared to central galaxies of similar mass (see Kereš et al. 2009; Simha et al. 2009). However, during the same time period, while central galaxies grow at a faster rate, their host haloes also accrete mass and ‘receive’ mergers of lower mass haloes. The balance between stellar mass growth and halo mass growth leads to a similar \( M_{H_1} \)–\( M_H \) relation at \( z = 0 \) for central and satellite systems, in both simulations.

Note that in Fig. 2 as well as in the remainder of this section, we only consider galaxies that are above our resolution threshold at \( z = 0 \), with \( M_S > 64 m_{\text{SPH}} = 5.8 \times 10^9 M_\odot \). However, some satellite galaxies in dense environments experience mass loss between \( z_{\text{sat}} \) and \( z = 0 \), and, consequently, a fraction of satellites that are above the resolution threshold at \( z_{\text{sat}} \) is pushed below it by \( z = 0 \). We defer examination of this point to Section 4.

We assign galaxies to the halo/subhalo population via the monotonic mapping procedure described in Section 2. The right-hand column of Fig. 2 shows the distribution of the ratio of the stellar mass assigned to a halo to the stellar mass of the SPH galaxy that is actually located within it for the SPHnw simulation (top) and the SPHw simulation (bottom). In addition to the distribution for all galaxies (black solid curve), we show the distribution for satellite galaxies alone (red dotted curve). For most haloes, the SHAM assigned mass is close to the mass of the SPH galaxy located within it both the SPHnw and the SPHw simulations, although there are a small number of extreme outliers. We characterize the width
of the distributions by $\sigma_M$ such that 68 per cent of galaxies have $|R| = |\log M_A/M_R| < \sigma_M$, where $M_A$ is the assigned mass and $M_R$ is the SPH (real) mass. In the SPHnw simulation, the width of the distribution is $\sigma_M = 0.09$, while in the SPHw simulation $\sigma_M = 0.11$. In both simulations, the distribution of SHAM assigned masses for satellite galaxies is also centred on the SPH mass, but there is greater scatter than in the case of central galaxies. For satellite galaxies in the SPHnw simulation, the distribution of assigned mass is skewed such that the assigned mass is likely to be slightly higher than the SPH mass.

Fig. 3 presents the same distribution of $|\log M_A/M_R|$ at $z = 0.5$, 1 and 2, and compared to the $z = 0$ result from Fig. 2. There is good agreement between the SHAM assigned stellar masses and the SPH stellar masses at all four epochs. The fraction of satellite galaxies decreases with increasing redshift. Despite this, $\sigma_M$ shows a continuous trend of increasing with redshift, from 0.09 at $z = 0$ to 0.13 at $z = 2$ in the SPHnw simulation, and from 0.11 at $z = 0$ to 0.14 at $z = 2$ in the SPHw simulation. Although the total baryonic mass (not shown) is equally well correlated with halo mass at $z = 2$ and at $z = 0$, the higher mean gas fraction and the larger halo-to-halo scatter in gas fraction at high redshift leads to higher scatter in the halo mass–stellar mass relation.

While the relationship between halo mass and galaxy stellar mass is roughly monotonic in our SPH simulations, there is some scatter, with the strongest outliers arising from satellite galaxies. We examine the sources of this scatter in Fig. 4, both to understand the physical processes that give rise to it and to explore the possibility of adding a parameter that would sharpen the SHAM. We restrict
Figure 3. Probability distribution of the ratio of stellar mass assigned by SHAM to SPH galaxy mass in the SPHnw simulation (left) and SPHw simulation (right). The solid, dot–dashed, dashed and dotted curves represent redshifts, $z = 0, 0.5, 1$ and $2$, respectively.

Figure 4. Stellar mass of satellite galaxies versus parent halo mass just before $z_{\text{sat}}$. Each point represents a galaxy, and only galaxies that are in the top 5 per cent or bottom 5 per cent by stellar mass in each 0.25 decade wide halo mass bin (relative to the distribution of all galaxies) are shown. In panel (a), the circles, triangles, squares and crosses stand for different $z = 0$ host halo mass bins while in panels (b) and (c), they stand for bins of $z_{\text{sat}}$, the epoch of accretion of the satellite and $\Delta M/M$, the change in stellar mass since $z_{\text{sat}}$, respectively. Panels (d)–(f) are analogous to panels (a)–(c) but using the SPHw simulation.
ourselves to \( z = 0 \) as the satellite galaxy sample is largest at this epoch. Each panel plots satellite galaxy stellar mass as a function of \( M_{\text{H}} \) at \( z_{\text{sat}} \) with SPHnw in the top panels and SPHw in the bottom.

In panel (a), points are colour coded by the mass of the \( z = 0 \) halo (not subhalo) that hosts the satellite, and lines show the median relation between \( M_{\text{H}}(z = 0) \) and \( M_{\text{H}}(z_{\text{sat}}) \) in four bins of \( z = 0 \) halo mass. The median curves are nearly identical for the four bins, indicating that the typical \( M_{\text{H}}/M_{\text{H}} \) is at most minimally correlated with the final halo mass. The outlier points at low \( M_{\text{H}}/M_{\text{H}} \) are mostly galaxies that have experienced stellar mass loss. These outliers are found in all \( M_{\text{H}} \) bins; one should not read too much into the relative numbers of outlier points as the total number of satellites varies from bin to bin. We caution that this figure does not show galaxies that have lost enough mass to fall below the \( M_{\text{H}} = 64 \, m_{\text{SPI}} \) threshold by \( z = 0 \) (see Fig. 11).

In panel (b), points and lines are coded by satellite accretion epoch. Once again, there is little difference in the median relations and outliers is found in all the \( z_{\text{sat}} \) bins (except \( z_{\text{sat}} > 2 \)). Panel (c) divides galaxies into those that have lost stellar mass since \( z_{\text{sat}} \), those that have increased their stellar mass by less than 10 per cent since \( z_{\text{sat}} \) and those that have increased their stellar mass by more than 10 per cent. Not surprisingly, galaxies that have lost stellar mass have systematically lower \( M_{\text{H}}/M_{\text{H}} \) at \( z = 0 \). Median relations for the other two populations are similar. There are some galaxies that are low \( M_{\text{H}}/M_{\text{H}} \) outliers despite having gained mass since \( z_{\text{sat}} \), indicating that at least some of this outlier population comes from satellites that had anomalously low \( M_{\text{H}} \) at \( z_{\text{sat}} \). Panels (d)–(f) repeat this analysis for the SPHw simulation. The results are qualitatively similar, though the smaller number of galaxies in this simulation makes it difficult to draw firm conclusions.

### 3.2 Luminosity and star formation rate

So far, we have used SHAM to assign stellar masses to haloes and compare them to the stellar masses of SPH galaxies located in those haloes. However, when SHAM is applied to an observed galaxy population, it is often used to assign luminosities to simulated haloes to compare the luminosity dependence of the clustering properties of the simulated haloes to that of the observed galaxies.

The luminosity of a galaxy is correlated with its stellar mass, but is not reducible to a simple function of stellar mass because the stars formed at different times. We track the star formation histories of simulated galaxies in our SPH simulations, and use the stellar population synthesis package of Conroy, Gunn & White (2009) to compute their luminosities. We assume that stars are formed with a Chabrier initial mass function. We assume solar metallicity and do not consider dust extinction. Changing the initial mass function in the same way for all galaxies would alter the zero-point of the luminosity–stellar mass relation but would not be likely to add scatter, while including dust extinction would shift the mean relation and increase the scatter somewhat.

The left-hand column of Fig. 5 shows the \( r \)-band luminosity of SPH galaxies against halo mass in the SPHnw simulation (top) and the SPHw simulation (bottom). For galaxies that are the central objects of their parent haloes (henceforth central galaxies), the halo mass, \( M_{\text{H}} \), on the horizontal axis, is the mass of the SO halo in which the galaxy is located at \( z = 0 \). For satellite galaxies, it is the mass of the SO halo in which the galaxy is located at \( z_{\text{sat}} \), the last output epoch before its parent halo fell into the virial radius of a more massive halo. The points show individual galaxies that are either in the top 5 per cent or bottom 5 per cent by luminosity in each halo mass bin. Most of the low-luminosity outliers are satellite galaxies. This is primarily because satellite galaxies typically have an older stellar population than central galaxies of similar stellar mass and are, therefore, less luminous. A secondary factor, of less importance, is the difference between central and satellite galaxies in the stellar mass–halo mass relation shown in Fig. 2.

To assign luminosities to our haloes, we rank order haloes by mass and assign simulated galaxies rank ordered by \( r \)-band luminosity to them, as done previously for stellar mass. The right-hand column of Fig. 5 shows the distribution of the ratio of the SHAM assigned luminosity of a halo to the luminosity of the SPH galaxy within it in the SPHnw simulation (top right) and the SPHw simulation (bottom right). In both simulations, the distribution of SHAM assigned luminosities is centred on the SPH luminosity. In analogy with the previous subsection, we define \( \sigma_L \), such that 68 per cent of galaxies have \( |R| = |\log L_\text{SHAM}/L_\text{SPH}| < \sigma_L \), where \( L_\text{SHAM} \) is the assigned luminosity and \( L_\text{SPH} \) is the SPH (real) luminosity. In the SPHnw as well as the SPHw simulation, \( \sigma_L = 0.15 \), which is greater than the corresponding \( \sigma_M \) (see Fig. 2). For satellite galaxies, the assigned luminosity is systematically higher than the SPH luminosity because of the stellar population differences, the offset being 0.08 dex in the SPHnw simulation but only 0.02 dex in the SPHw simulation.

The clustering properties of \( z \geq 2 \) galaxies are studied observationally using the Lyman break technique, in which high-redshift star-forming galaxies are identified by optical photometry alone using their redshifted rest-frame UV radiation (Steidel et al. 1996, 1999, 2003). Conroy et al. (2006) find good agreement between the clustering of Lyman break galaxies (LBGs) and the clustering of haloes in their \( N \)-body simulation when they use SHAM to match haloes to galaxies. The relationship between LBGs and their host haloes is important in understanding the properties of LBGs, in particular whether they are a quiescent star-forming population (Coles et al. 1998; Mo, Mao & White 1999; Giavalisco & Dickinson 2001) or a merger-driven starburst population (Somerville, Primack & Faber 2001; Scannapieco & Thacker 2003). These applications raise the question of whether SHAM can be reliably applied to the rest-frame UV luminosities of high-redshift galaxies, which depend mainly on their SFRs rather than their stellar masses.

In our simulations, the SFRs and stellar masses of high-redshift galaxies are well correlated (see e.g. fig. 7 of Davé et al. 2010). In the left-hand panels of Fig. 6, we directly investigate the relationship between the \( z = 2 \) SFRs of our simulated galaxies and the properties of their parent haloes, in the same format used previously for stellar mass and \( r \)-band luminosity at \( z = 0 \). To calculate the SFR of an SPH galaxy, we sum over the SFRs of all its gas particles computed according to the formula in Katz et al. (1996). In contrast to \( z = 0 \), where there is a significant passive, non-star-forming population, only \( \sim 0.1 \) per cent of galaxies have no star formation at \( z = 2 \), and these are excluded from the analysis. In the absence of dust extinction, the instantaneous SFR should be a good indicator of rest-frame UV luminosity, since the latter is dominated by the output of young, short-lived stars. We caution, however, that dust extinction corrections for LBG UV luminosities are typically factors of several (e.g. Steidel et al. 2003), and a scatter in extinction at fixed SFR could add significant scatter to the relation between \( M_{\text{H}} \) and UV luminosity. Since colours provide an indication of extinction, the most effective strategy for observational analysis is probably to apply colour-based extinction corrections before SHAM, so that only the errors in the corrections add scatter.

Fig. 6 shows that the relation between intrinsic SFR and \( M_{\text{H}} \) at \( z = 2 \) is nearly as tight as the relation between \( r \)-band luminosity and \( M_{\text{H}} \) at \( z = 0 \). We can therefore assign SFRs to SHAM galaxies via the same monotonic matching used previously for stellar mass and...
Figure 5. Left: R-band luminosity at \( z = 0 \) versus halo mass in the SPHnw simulation (top) and the SPHw simulation (bottom). Each point represents a galaxy in the SPH simulation. For central galaxies, shown as black points, the halo mass is the \( z = 0 \) mass of the host halo, while for satellite galaxies, shown as red points, the halo mass is the mass of the parent halo just before \( z_{\text{sat}} \). Only galaxies that are in the top 5 per cent or bottom 5 per cent by R-band luminosity in each 0.25 decade wide halo mass bin are shown. The solid and dotted curves show the median R-band luminosity in each halo mass bin for all galaxies and satellite galaxies, respectively. Right: probability distribution of the ratio of R-band luminosity assigned by SHAM to R-band luminosity of SPH galaxies computed using a stellar population synthesis code in the SPHnw simulation (top) and SPHw simulation (bottom), with the solid curve standing for all galaxies and the dotted curve for satellite galaxies.

\( R \)-band luminosity, with the results shown in the right-hand panels. The 68 per cent scatter of \( \log \frac{\text{SFR}_A}{\text{SFR}_R} \) is \( \sigma_S = 0.12 \) and 0.16 for the SPHnw and SPHw simulations, respectively, compared to the \( R \)-band luminosity scatters of \( \sigma_L = 0.15 \) for both simulations at \( z = 0 \). The median relations for central and satellite galaxies are nearly the same, in both cases, with offsets that are small compared to the intrinsic scatter. However, in the no-wind simulation the outliers at low SFR are preferentially satellites, a result of the gradual shutdown of gas accretion after galaxies become satellites in larger haloes (Kereš et al. 2009; Simha et al. 2009). The \( \log \frac{\text{SFR}_A}{\text{SFR}_R} \) histogram is, therefore, skewed towards underestimated SFRs for satellites, though this remains a small effect. For the wind simulation, there are many fewer galaxies above our 64 \( m_{\text{SPH}} \) stellar mass threshold, and SFRs at \( M_H < 10^{12.4} M_\odot \) are suppressed by the momentum-driven outflows. In this simulation, there are more satellite outliers at high SFR, and the (noisy) histogram of \( \log \frac{\text{SFR}_A}{\text{SFR}_R} \) for satellite galaxies shows an overall shift towards underestimated SFR. For central galaxies, gas fractions in the SPHw simulation are systematically higher than those in the SPHnw simulation, and scatter in these central galaxy gas fractions (contributing scatter in SFR) is the largest factor driving higher \( \sigma_S \) for the SPHw case.

Fig. 6 suggests that SHAM should be a reliable tool for assigning SFRs to simulated halo populations at high redshift (given an observed SFR distribution), or for inferring the halo masses associated with observed LBGs. We caution that our simulations do not produce a substantial population of passive galaxies at these redshifts.
redshifts, either central or satellite, and that physical mechanisms that strongly suppress star formation at high stellar mass could reduce the accuracy of SHAM assignment.

4 SHAM IN THE DM SIMULATION

So far we have applied SHAM to the halo and subhalo hosts of galaxies in the SPH simulations, where the condensed baryons may improve the survival of DM substructures. We now turn to SHAM as it is traditionally applied by using the DM simulation, which starts from the same initial conditions but includes no baryons. Independent SO haloes above the resolution limit have similar locations and masses in our SPHnw, SPHw and DM simulations. We identify DM substructures within these SO haloes using the ADAPTAHOP code as described earlier. While there is reasonable agreement in the number and abundance of subhaloes between the two SPH simulations and the DM simulation, there are positional differences. For each subhalo above our resolution threshold of 64 particles, we trace its high-redshift progenitors up to the epoch when it was an independent halo, i.e. up to $z_{sat}$.

The solid curve in each panel of Fig. 7 shows $\langle N(M) \rangle$, the mean number of SPH galaxies per halo above a given stellar mass threshold in each halo mass bin in the SPHnw simulation. The four panels correspond to different stellar mass thresholds, and the mean space density of galaxies above these thresholds ranges from 0.004 to $0.06 h^3$ Mpc$^{-3}$. The dashed curve shows the mean number of galaxies per halo when haloes in the DM simulation are populated with galaxies using SHAM as described in Section 3, with a monotonic relation between stellar mass and $M_{H}$ (equation 2). For comparison, the dotted curves show results of a model (SHAMz0) in which we assume a monotonic relationship between galaxy stellar mass and $z = 0$ subhalo mass, rather than $z_{sat}$ subhalo mass. Note that the procedures of Vale & Ostriker (2004) and Weinberg et al. (2008) differ from SHAMz0 because they apply a global mass-loss correction to subhalo masses, though they do not consider post-$z_{sat}$ mass loss on an object-by-object basis.
Testing SHAM in SPH simulations

Figure 7. Mean number of galaxies per halo versus halo mass. The solid curve represents the SPHnw simulation, while the dashed and dotted curves are results from populating haloes in our $N$-body simulation with galaxies using SHAM with subhalo masses at $z_{\text{sat}}$ and at $z=0$ (SHAMz0), respectively. Each panel stands for a different galaxy stellar mass threshold, for which the corresponding space density of galaxies is also indicated. Curves are truncated when the mean occupation of haloes in the next lower (0.1 dex) mass bin falls below 0.03.

In low-mass haloes, where the satellite fraction is low, both the SHAMz0 and SHAM model predictions for the average number of galaxies per halo are in good agreement with the SPH simulation. At high halo mass, however, SHAMz0 underpredicts the number of galaxies per halo because it does not account for subhalo mass loss. The stellar masses assigned to subhaloes are too low, and galaxies that should be above a stellar mass threshold instead fall below it. SHAM, on the other hand, remains in good agreement with the SPH simulation over a wide range of halo masses and galaxy stellar mass thresholds.

Fig. 8 is the analogue of Fig. 7, but using the SPHw simulation instead of the SPHnw simulation. As winds reduce stellar mass, the number density of galaxies above a given stellar mass threshold is lower than in the SPHnw simulation. Nonetheless, the trends for the SPHw simulation discussed above still hold. Abundance matching using the $z=0$ mass of subhaloes underpredicts the number of galaxies in massive haloes. Although, SHAM (using the $z_{\text{sat}}$ mass of subhaloes) provides a better fit to the SPH galaxy sample, it overpredicts the number of galaxies in massive haloes to a larger degree than in the SPHnw simulation. The one-halo term of the galaxy correlation function depends on the mean pair number $\langle N(N-1) \rangle$. We have checked that the standard prescription (Kravtsov et al. 2004; Zheng et al. 2005) that satellite galaxies follow Poisson statistics ($\langle N_{\text{sat}}(N_{\text{sat}}-1) \rangle = \langle N_{\text{sat}} \rangle^2$ describes both simulations well, allowing computation of $\langle N(N-1) \rangle$ from $\langle N(M) \rangle$.

Fig. 9 compares the halo occupations of SPH galaxies to the SHAM and SHAMz0 populations in each of the 30 most massive haloes at $z=0$, for galaxies above the $64 m_{\text{SPH}} = 5.8 \times 10^9 M_\odot$ threshold. As already seen in Figs 7 and 8, SHAMz0 predicts too few galaxies in massive haloes in both simulations. In the SPHnw simulation, the agreement between the number of SPH galaxies in each halo and the number of galaxies assigned to the halo by SHAM is remarkably good, indicating that SHAM is not just reproducing the typical number of galaxies for a given halo mass, but is also capturing the variation in galaxy number at a given halo mass. In the SPHw simulation, however, SHAM more noticeably overpredicts
the number of galaxies in the most massive haloes, and it does not track the variation in galaxy number at a given halo mass as well as in the SPHnw simulation.

Galaxy clustering depends on halo occupation statistics like those shown in Figs 7–9, and on small scales it also depends on the radial profile of satellites in massive haloes. Fig. 10 compares the radial number density profile of SPH galaxies around the central galaxy of the halo to the radial number density profile of galaxies assigned to haloes by SHAM and SHAMz0. The left-hand panels correspond to the single most massive halo in the simulations, with $M_{\text{halo}} = 4 \times 10^{14} \, M_{\odot}$, while the right-hand panels correspond to the three haloes with $10^{14} < M_{\text{halo}} < 3 \times 10^{14} \, M_{\odot}$. In all cases, the slopes of the radial density profiles are similar for SPH galaxies and for subhaloes populated by SHAM and by SHAMz0, but the normalizations and the inner truncations (indicated by where the curves stop) differ. The normalization differences correspond to the differences in halo occupation at high $M_{\text{halo}}$, seen in Figs 8 and 9. SHAMz0 predictions are always suppressed relative to SPH galaxies because they neglect subhalo mass loss. Tidal stripping is more severe for subhaloes near the halo centre, exacerbating this effect and causing truncation of the SHAMz0 profiles at larger radii compared to SPH galaxies. Weinberg et al. (2008) found a similar truncation effect even when including a global correction for subhalo mass loss.

In contrast to SHAMz0, SHAM overpredicts the number of galaxies in high-mass haloes, by a small factor in the SPHnw simulation, and by 0.1–0.3 dex in the SPHw simulation. This overprediction is a consequence of neglecting stellar mass loss that occurs after $z_{\text{cut}}$ in the SPH simulation, which is not accounted for in the SHAM recipe. The stellar mass loss is more severe for satellites close to the halo centre, so in this case the SPH profiles truncate at larger radii than the SHAM profiles.

The overprediction of satellite numbers by SHAM may seem surprising in light of the good agreement between assigned and true stellar masses seen in Fig. 2, with distributions that peak at $\log M_{\text{ext}}/M_{\odot} = 0$ and have only mild asymmetry. However, the relations in Fig. 2 (and the versions divided by halo mass in Fig. 4) only include galaxies that remain above the 64 $m_{\text{SPH}}$ stellar mass threshold at $z = 0$. In high-mass haloes, a significant fraction of

Figure 8. Mean number of galaxies per halo versus halo mass. This figure is analogous to Fig. 7 but using the SPHw simulation, which includes momentum-driven winds.

© 2012 The Authors, MNRAS 423, 3458–3473
Monthly Notices of the Royal Astronomical Society © 2012 RAS
For the SPHnw simulation, agreement between the SHAM and SPH–galaxy correlation functions is excellent, as one would expect from the close agreement of halo occupations and radial profiles seen in earlier figures. The largest discrepancies are ~10 per cent for both 2 < r ≤ 10 h⁻¹ Mpc and r ≤ 2 h⁻¹ Mpc. For SPHw, the agreement in the two-halo regime is still very good, with a maximum discrepancy of ~15 per cent for 2 < r ≤ 10 h⁻¹ Mpc. However, the overpopulation of high-mass haloes leads to substantial over-prediction of ξ(r) in the one-halo regime, by up to a factor of ~2.5. This discrepancy arises from the severe stellar mass loss that affects a small but not negligible fraction of satellites in high-mass haloes and is not captured by the SHAM recipe. The impact on three-point or higher order correlation functions, which more strongly weight the single-halo occupations at small scales, would be more severe.

Panels (c) and (d) of Fig. 11 are the analogues of panels (a) and (b), respectively, but only show galaxies whose parent haloes are more massive than 10¹² M⊙ (at z = 0 for central galaxies and z_sat for satellite galaxies). This corresponds to a space density of galaxies of 0.007 h³ Mpc⁻³ and a stellar mass threshold of 7.3 × 10¹⁰ M⊙ in the no winds simulation and 2.8 × 10¹⁰ M⊙ in the winds simulation. For the SPHnw simulation, the trends discussed above still hold, agreement between the SHAM and SPH–galaxy correlation functions is better than ~18 per cent on all scales. For the SPHw simulation, in contrast with panel (b), the agreement between the SHAM predicted and SPH–galaxy correlation functions is better than ~20 per cent on all scales except r ≤ 0.3 h⁻¹ Mpc, where the discrepancy rises to ~50 per cent. Compared to galaxies hosted by low-mass subhaloes, a substantially smaller fraction of galaxies hosted by more massive subhaloes undergo severe stellar mass loss. Furthermore, for subhaloes with M₁(zi) ≥ 10¹² M⊙, the fraction of satellites that survive up to z = 0 is only marginally lower in the SPHw simulation compared to the SPHnw simulation.

5 CONCLUSIONS

Ever since the identification of substructures in N-body simulations (e.g. Ghigna et al. 1998; Klypin et al. 1999; Moore et al. 1999; Springel et al. 2001), there have been efforts to associate them with...
Figure 10. Radial number density profile of SPH galaxies (black solid curve) and haloes in the N-body simulation populated with galaxies using SHAM (blue dashed curve) and SHAMz0 (red dot–dashed curve), for the mass threshold of $M_S = 5.8 \times 10^9 M_\odot$. The top two panels are for two halo mass bins in the SPHnw simulation, while the bottom two panels are for two halo mass bins in the SPHw simulation. The curves stop when the only interior galaxy is the central galaxy. Also shown are radial number density profiles of SPH galaxies (grey solid curve) and SHAM (purple dashed curve) selected haloes using a higher threshold of $M_S = 7.3 \times 10^{10} M_\odot$ in the no winds simulation and $M_S = 2.8 \times 10^{10} M_\odot$ in the winds simulation.

galaxies (e.g. Colín et al. 1999; Kravtsov et al. 2004; Weinberg et al. 2008). SHAM has had impressive empirical success, reproducing observed galaxy clustering over a wide range of luminosity and redshift and correctly diagnosing (via cluster mass-to-light ratios) the overestimated matter clustering amplitude ($\sigma_8 \approx 0.9$ versus $\sigma_8 \approx 0.8$) in 1-year Wilkinson Microwave Anisotropy Probe (WMAP1)-era cosmological models (Conroy et al. 2006; Vale & Ostriker 2006; Guo et al. 2010; Moster et al. 2010). Our investigation provides the first full-scale test of SHAM against hydrodynamic cosmological simulations, where the correct identification between galaxies and subhaloes is known a priori, including both a simulation with minimal feedback (SPHnw) and a simulation with momentum-driven winds that better reproduces the observed galaxy stellar mass function (Oppenheimer et al. 2010). This investigation yields physical insight into the galaxy formation process in these simulations, and it demonstrates the strengths and potential limitations of SHAM as a tool for interpreting observed galaxy clustering, significantly extending earlier studies by Nagai & Kravtsov (2005) and Weinberg et al. (2008).

When we consider galaxies above our adopted stellar mass threshold at $z = 0$, $M_S \geq 64 m_{SPH} = 5.8 \times 10^9 M_\odot$, we find a tight correlation between galaxy stellar mass and the mass of the parent halo or subhalo. For central galaxies, we consider the full mass of the spherical overdensity at $z = 0$, while for satellite galaxies we use the mass of the parent halo just before the epoch $z_{sat}$ when it first becomes a satellite. Importantly, the median relation and scatter between $M_S$ and $M_H$ are similar for central galaxies and satellite galaxies, in both simulations, and the outlier fraction for satellite galaxies is only modestly higher, mainly because of stellar mass loss in some satellites after $z_{sat}$. As a result, SHAM assignment of stellar masses is remarkably effective in both simulations. The 68 per cent scatter in $R = \log M_A/M_R$, where $M_A$ is the assigned

galaxies.
Figure 11. Two-point correlation function of SPH galaxies (solid curve) and haloes and subhaloes in the $N$-body simulation populated using SHAM (dashed curve) and SHAMz0 (dot–dashed curve). Panels (a) and (b) show all galaxies above the stellar mass resolution threshold ($5.8 \times 10^9 \, M_\odot$) in the SPHnw (no winds) and SPHw (winds) simulations, respectively. Panels (c) and (d) show galaxies above a $10^{12} \, M_\odot$ halo mass threshold (at $z = 0$ for central galaxies or $z_{\text{sat}}$ for satellite galaxies). Lower windows of each panel show the ratio of the SHAM correlation function to the SPH galaxy correlation function.

stellar mass and $M_R$ the real (simulation) stellar mass, is $\sigma_M = 0.09$ dex for SPHnw and 0.11 dex for SPHw. The distribution of $R$ is only mildly asymmetric, with a small extended tail of outliers. Similar results, with slightly increased scatter, hold at $z = 0.5$, 1 and 2. We find no clear correlation between residuals from the $M_S$--$M_H$ relation and the satellite epoch at $z_{\text{sat}}$ or the parent mass of the $z = 0$ halo, so our tests do not suggest a way to further tighten SHAM by considering additional properties.

Using $R$-band luminosity in place of stellar mass yields similar results, but with a larger population of outliers among satellites caused by their systematically older stellar populations and thus higher $M_S/L$ ratios. SHAM should, therefore, be applied to stellar masses (estimated from luminosity and colour or spectral energy distribution) when possible, or to luminosity in redder bands that more faithfully trace stellar mass. At $z = 2$, instantaneous SFRs, which should be a good proxy for observed-frame optical luminosities, are well correlated with $M_H$, with similar correlations for central and satellite galaxies. SHAM assignment of SFRs at this redshift is quite effective with a scatter in log SFR$_i$/SFR$_R$ of 0.12 dex in SPHnw and 0.16 dex in SPHw. This result reinforces empirical evidence, based on galaxy clustering data, that SHAM is an effective tool for modelling LBGs at high redshift.

SHAM is traditionally applied to $N$-body rather than SPH simulations (or to analytic descriptions calibrated on $N$-body). To test this standard form of SHAM, we have applied it to the ADAPTAHOP subhalo population of a pure DM simulation started from the same initial conditions as the SPH simulations. Because subhalo positions within parent haloes shift between SPH and DM (Weinberg et al. 2008), we have focused our comparison on halo occupation statistics and radial profiles, which together determine many properties of observable galaxy clustering, and on the real-space two-point correlation function.

Using the galaxy stellar mass function of the SPHnw or SPHw simulation as input, SHAM (applied to the DM simulation) does quite well in reproducing the corresponding mean halo occupation, $\langle N(M_{\text{halo}}) \rangle$, and the slope of the galaxy radial profile in high-mass haloes. In SPHnw, but not SPHw, SHAM traces the individual halo-to-halo variations in galaxy number at similar halo mass. Use
of subhalo masses at $z_{\text{sat}}$ rather than $z = 0$ makes a critical difference; if we use $z = 0$ subhalo masses, the galaxy occupation in high-mass haloes is systematically depressed because subhaloes lose mass by tidal stripping after becoming satellites. Traditional SHAM (using $z_{\text{sat}}$ masses), by contrast, tends to overpredict the galaxy numbers in high-mass haloes. We trace this discrepancy to the small but not negligible population of satellite galaxies that suffer severe stellar mass loss in the SPH simulations, so that they move from above our resolution threshold at $z_{\text{sat}}$ to below it at $z = 0$. The subhaloes themselves remain identifiable in the DM simulation, and SHAM populates them with galaxies that lie on the main $M_{\text{vir}}-M_{\text{halo}}$ relation. The effect is more significant in the SPHw simulation, whose galaxies are apparently more vulnerable to severe mass loss because of their shallower baryonic potential wells. The mean occupation of high-mass haloes in SPHw is overpredicted by 0.1–0.3 dex, while in SPHnw the offset is less than 0.1 dex. For SPHnw, the SHAM-predicted correlation function agrees with that of SPH galaxies to better than 10 per cent at all scales 0.05 < $r$ < 10 $h^{-1}$ Mpc. For SPHw, SHAM predicts the correlation function to 15 per cent at $r > 2$ $h^{-1}$ Mpc but it overpredicts by a factor of $\sim 2.5$ at $r < 0.5$ $h^{-1}$ Mpc because of the overpopulation of massive haloes. However, if the stellar mass threshold is increased to $2.7 \times 10^{10} M_{\odot}$, then the SHAM predicted two-point correlation function agrees with that of SPH galaxies to better than 20 per cent on all scales except $r < 0.3$ $h^{-1}$ Mpc, where the discrepancy widens to 50 per cent. This improvement is achieved because galaxies that suffer severe mass loss are typically low stellar mass galaxies in low-mass subhaloes ($M_{\text{halo}} \leq 10^{12} M_{\odot}$). While we would expect low-mass galaxies in low-mass subhaloes to be more susceptible to severe mass loss, further investigation with higher resolution simulations is required to understand the extent to which this is a numerical artefact caused by proximity of the mass to the threshold above which we are reliably able to resolve galaxies. For galaxies with luminosity $L > L_{\ast}$, our results suggest that SHAM will modestly (by $\sim$10–30 per cent) overpredict the two-point correlation function on small scales and yield accurate clustering predictions on larger scales.

In both of our SPH simulations, halo mass (at $z = 0$ for central galaxies or $z_{\text{sat}}$ for satellite galaxies) is the primary determinant of galaxy stellar mass, luminosity and (at high redshift) SFR. The most significant secondary factor, and thus the most significant limitation on the accuracy of SHAM, is the small fraction of satellite galaxies that suffer severe stellar mass loss even though their host subhaloes survive. The sensitivity to feedback, indicated by the difference between our SPHnw and SPHw simulations, motivates further investigation of galaxy mass loss in high-mass haloes for a wider range of feedback prescriptions.

For galaxies with $L \geq L_{\ast}$, the SPHw simulation predicts excessive galaxy masses and excessive late-time star formation, indicating the need for an additional physical mechanism such as active galactic nucleus (AGN) feedback. The accuracy of SHAM for high-luminosity galaxies will depend on how tightly correlated such feedback is with halo mass. Overall, however, the strong role of DM in governing galaxy formation makes SHAM a powerful technique for making realistic artificial galaxy catalogues and for interpreting the observed distribution of galaxies.

ACKNOWLEDGMENTS

We thank Stephane Colombi, Charlie Conroy and Andrey Kravtsov for providing the ADAPT code and for assistance in using it. We thank Charlie Conroy for providing the stellar population code used to compute galaxy luminosities. This work was supported by NSF grant AST0707985 and NASA ATP grant NNX10A956.

REFERENCES

Aubert D., Pichon C., Colombi S., 2004, MNRAS, 352, 376
Barnes J., Hut P., 1986, Nat, 324, 446
Berlind A. A. et al., 2003, ApJ, 593, 1
Coles P., Lucchin F., Matarrese S., Moscardini L., 1998, MNRAS, 300, 183
Colin P., Klypin A. A., Kravtsov A. V., Khokhlov A. M., 1999, ApJ, 523, 32
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Conroy C., Gunn J. E., White, M., 2009, ApJ, 699, 486
Davé R., Finlator K., Oppenheimer B. D., Fardal M., Katz N., Kereš D., Weinberg D. H., 2010, MNRAS, 404, 1355
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Evrard A. E., Summers F. J., Davis M., 1994, ApJ, 422, 11
Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
Gelb J. M., Bertschinger E., 1994, ApJ, 436, 467
Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 1998, MNRAS, 300, 146
Giavalisco M., Dickinson M., 2001, ApJ, 550, 177
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Haardt F., Madau P., 2001, in Neumann D. M., Van J. T. T., eds, Clusters of Galaxies of the High Redshift Universe Observed in X-rays preprint (astro-ph/0106018)
Hernquist L., 1987, ApJS, 64, 715
Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles. McGraw-Hill, New York
Katz N., 1992, PASP, 104, 852
Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19
Kazantzidis S., Mayer L., Mastropietro C., Diemand J., Stadel J., Moore B., 2004, ApJ, 608, 663
Kennicutt R. C., Jr, 1998, ApJ, 498, 541
Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, MNRAS, 395, 160
Kitayama T., Suto Y., 1996, ApJ, 469, 480
Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Knebe A. et al., 2011, MNRAS, 415, 2293
Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
Larson D. et al., 2011, ApJS, 192, 16
Mo H. J., Mao S., White S. D. M., 1999, MNRAS, 304, 175
Hoare D., Quinn T., Governato F., Stadel J., Lake G., 1999, MNRAS, 310, 1147
Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903
Murali C., Katz N., Hernquist L., Weinberg D. H., Davé R., 2002, ApJ, 571, 1
Nagai D., Kravtsov A. V., 2005, ApJ, 618, 557
Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265
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
Reid B. A. et al., 2010, MNRAS, 404, 60
Scannapieco E., Thacker R. J., 2003, ApJ, 590, L69
Schmidt M., 1959, ApJ, 129, 243
Simha V., Weinberg D. H., Davé R., Gnedin O. Y., Katz N., Kereš D., 2009, MNRAS, 399, 650
Somerville R. S., Primack J. R., Faber S. M., 2001, MNRAS, 320, 504
Springel V., 2005, MNRAS, 364, 1105
Springel V., Hernquist L., 2002, MNRAS, 333, 649
Springel V., Hernquist L., 2003, MNRAS, 339, 289

© 2012 The Authors, MNRAS 423, 3458–3473
Monthly Notices of the Royal Astronomical Society © 2012 RAS
Testing SHAM in SPH simulations

Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
Steidel C. C., Giavalisco M., Dickinson M., Adelberger K. L., 1996, AJ, 112, 352
Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1
Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2003, ApJ, 592, 728
Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, ApJ, 742, 16

Vale A., Ostriker J. P., 2004, MNRAS, 353, 189
Vale A., Ostriker J. P., 2006, MNRAS, 371, 1173
Weinberg D. H., Colombi S., Davé R., Katz N., 2008, ApJ, 678, 6
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Zheng Z. et al., 2005, ApJ, 633, 791
Zu Y., Weinberg D. H., Davé R., Fardal M., Katz N., Kereš D., Oppenheimer B. D., 2011, MNRAS, 412, 1059

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