The pre-processing of subhaloes in SDSS groups and clusters

Annie Hou, Laura C. Parker and William E. Harris

Department of Physics and Astronomy, McMaster University, Hamilton ON L8S 4M1, Canada

Accepted 2014 April 26. Received 2014 April 23; in original form 2013 July 4

ABSTRACT

We investigate pre-processing using the observed quenched fraction of group and cluster galaxies in the Yang et al. Sloan Digital Sky Survey (SDSS)-seventh data release (DR7) group catalogue in the redshift range of $0.01 < z < 0.045$. We categorize group galaxies as virialized, infall or backsplash and we apply a combination of the Dressler–Shectman statistic and group member velocities to identify subhaloes. On average, the fraction of galaxies that reside in subhaloes is a function of host halo mass, where more massive systems have a higher fraction of subhalo galaxies both in the overall galaxy and infall populations. Additionally, we find that within the range $2 \lesssim r_{200} < 3$ the quiescent fraction is higher in the subhalo population with respect to both the field and non-subhalo populations. At these large radii ($2 \lesssim r_{200} < 3$), the majority of galaxies ($\sim 80$ per cent) belong to the infall population and therefore, we attribute the enhanced quenching to infalling subhalo galaxies, indicating that pre-processing has occurred in the subhalo population. We conclude that pre-processing plays a significant role in the observed quiescent fraction, but only for the most massive ($M_{\text{halo}} > 10^{14.5} \, M_\odot$) systems in our sample.

Key words: galaxies: clusters: general – galaxies: formation – galaxies: groups: general.

1 INTRODUCTION

Observational studies of rich galaxy clusters have shown that most of the members are red early-type galaxies with little or no on-going star formation (Oemler 1974; Dressler 1980; Blanton et al. 2003; Balogh et al. 2004; Baldry et al. 2006). While a high fraction of quiescent (i.e. not actively star-forming) galaxies have been observed in rich groups and clusters (Kauffmann et al. 2004; Wilman et al. 2005; Peng et al. 2010; McGee et al. 2011; Muzzin et al. 2012), recent results from observations and simulations (both numerical and semi-analytic) indicate that star formation quenching actually begins in low-mass haloes with $M_{\text{halo}} \sim 10^{13} \, M_\odot$ (McGee et al. 2009; Balogh & McGee 2010; George et al. 2011; De Lucia et al. 2012; Wetzel, Tinker & Conroy 2012). Additionally, there is evidence that some cluster galaxies had their star formation quenched in groups with $M_{\text{halo}} \geq 10^{13} \, M_\odot$ prior to accretion on to the more massive cluster environment, a process often referred to as pre-processing (Zabludoff & Mulchaey 1998a; Kawata & Mulchaey 2008; Berrier et al. 2009; McGee et al. 2009; De Lucia et al. 2012).

While quenching has been shown to occur in low-mass haloes, the significance of pre-processing is still a subject of debate. Using $N$-body simulations, Berrier et al. (2009) found that 70 per cent of their cluster ($10^{12} < M_{\text{halo}} < 10^{14.5} \, M_\odot$) galaxies fell in directly from the field, while only $\sim 10$ per cent fell in as members of groups. Size haloes with $M_{\text{halo}} \geq 10^{13} \, M_\odot$. Based on these results, Berrier et al. (2009) concluded that pre-processing did not significantly contribute to the quenched fractions observed in present-day clusters. In contrast, both McGee et al. (2009) and De Lucia et al. (2012) used semi-analytic models (SAMs) to show that $\sim 25$–$45$ per cent of their simulated cluster galaxies fell in as members of systems with $M_{\text{halo}} \geq 10^{13} \, M_\odot$, where the range depends on the mass of the galaxy and the mass of the host cluster. It should be noted that according to De Lucia et al. (2012), part of the discrepancy between the results of Berrier et al. (2009) and McGee et al. (2009) arises from differing definitions of ’satellite’, with the former computing fractions based on the time when a galaxy first becomes a satellite of any halo and the latter when a galaxy becomes a satellite of the final or present-day group or cluster. With the former definition, De Lucia et al. (2012) find that their results are not inconsistent with those of Berrier et al. (2009). A similar analysis was carried out using $N$-body hydrodynamical simulations by Bahé et al. (2013). These authors found that $\sim 15$–$60$ per cent of galaxies in host haloes in the mass range of $10^{13.5} < M_{\text{halo}} < 10^{15.2} \, M_\odot$ had been pre-processed where the amount of pre-processing scaled with halo mass; massive haloes had a higher fraction of pre-processed galaxies (Bahé et al. 2013).

Thus, the results of some SAMs (e.g. McGee et al. 2009; De Lucia et al. 2012) and numerical simulations (e.g. Bahé et al. 2013) predict that pre-processing can play an important role in quenching star formation, especially in massive clusters. If the simulation predictions of significant pre-processing in groups and clusters are correct then it should be possible to observe pre-processing by looking at the populations of galaxies in different environments. The aim

*E-mail: anniehou.how@gmail.com

© 2014 The Authors
Published by Oxford University Press on behalf of the Royal Astronomical Society
of this paper is to investigate the significance of pre-processing in a statistical sample of observed groups and clusters.

Pre-processing can be investigated by studying the properties of infalling subhalo galaxies, where a subhalo is defined as a collection of galaxies that reside in a small halo embedded within a larger parent halo. Subhaloes can be identified by performing substructure analysis with the Dressler–Shectman (DS) Test (Dressler & Shectman 1988), which can detect galaxies with kinematic properties that deviate from those of the host halo. It should be noted that this method of identifying subhaloes differs from those used in numerical simulations. In particular, our observational definition of subhaloes is based on identification of kinematically distinct galaxies and does not require the galaxies within the subhalo to be gravitationally bound to one another, which is usually the case for subhaloes identified in simulations. Subhaloes, detected via the DS Test, are preferentially found on the group or cluster outskirts (West & Bothun 1990; Zabludoff & Mulchaey 1998b; Hou et al. 2012; Dressler et al. 2013) and the usual assumption is that these systems are infalling. However, numerical simulations have shown that a large fraction of galaxies beyond the virial radius, and out to ~2.5 virial radii, have already passed through the group or cluster core (i.e. backsplash galaxies: Balogh, Navarro & Morris 2000; Mamon et al. 2004; Gill, Knebe & Gibson 2005; Mahajan, Mamon & Raychaudhury 2011; Pimbblet 2011; Bahé et al. 2013; Oman, Hudson & Behroozi 2013). Backsplash galaxies may have experienced star formation quenching due to more massive group – or cluster-related processes, and it has been suggested that much of the environmental quenching beyond the virial radius (out to ~2.5 virial radii) is most likely due to the presence of a backsplash population (Wetzel et al. 2014). In contrast, the infall population typically refers to galaxies that are infalling on to the host system for the first time. Thus, any observed enhanced quenching must be a result of a transformation that occurred prior to accretion on to the host halo.

Currently, most methods of distinguishing between the virialized, infall and backsplash populations in observed groups and clusters are based on the results of simulated systems. These classification schemes typically involve examining Δc(z)/σ distributions (Gill et al. 2005; Pimbblet 2011) or dividing the Δc(z)/σ−2D plane into regions occupied by virialized, infall and backsplash galaxies (Mahajan et al. 2011; Oman et al. 2013). Although each population resides in a distinct region in the full phase-space of simulated clusters, projection effects can distort these clear divisions and there is often contamination between the observed populations (to be discussed in more detail in Section 4.4). In addition to differences in their phase-space locations, infall and backsplash galaxies should also have subtle differences in their stellar mass distributions. As a result of tidal disruption, backsplash galaxies should be on average less massive than infalling galaxies at the same radius (Gill et al. 2005), and galaxies infalling in subhaloes will typically be more massive than individual infalling galaxies (McGee et al. 2009). Thus, in order to better probe pre-processing and environmental effects on galaxy evolution, it is important to examine the properties of virialized, infall and backsplash galaxies as a function of stellar mass and over a wide range of masses.

Although it is well known that high-density environments, such as groups and clusters, show signs of enhanced star formation quenching with respect to the field (Kauffmann et al. 2004; Rines et al. 2005; Kimm et al. 2009; Wetzel et al. 2012; Woo et al. 2013), the processes that dominate this transformation are still debated. Comparing the properties of infalling and backsplash subhalo galaxies allows us to probe the relative importance of rich group- and cluster-related processes, which are observable in the backsplash population, to pre-processing in lower mass haloes, which can be observed in the infalling subhalo population.

In this paper, we use a well-studied Sloan Digital Sky Survey (SDSS) group catalogue to probe the properties of subhalo galaxies in groups and clusters in order to investigate the amount of pre-processing that occurs and to study the relative importance of the lower mass group environment in the evolution of galaxies. The paper is structured as follows: in Section 2, we present our group and galaxy sample and in Section 3, we discuss how we identify subhaloes. We compare the properties of the non-subhalo and subhalo populations, as well as compare the virialized, infall and backsplash subpopulations in Section 4. Finally, in Section 5, we discuss our results and present our conclusions in Section 6. Throughout this paper, we assume a ΛCDM cosmology with Ωm,0 = 0.31, ΩΛ,0 = 0.69 and H0 = 70 km s−1 Mpc−1.

2 DATA

Observational results have found correlations between host environments and galaxy properties; however, these correlations are more easily observed in low-mass galaxies (e.g. Peng et al. 2010; Carollo et al. 2013; Hou et al. 2013). Thus, to fully investigate the role of pre-processing in galaxy groups, we require a large sample of group and cluster galaxies that is complete down to low stellar masses (log10(M∗/M⊙) ∼ 9.5), where environmental trends are expected to be more significant. These requirements can be achieved with the Yang et al. (2007) SDSS group catalogue.

2.1 The SDSS-DR7 Galaxy catalogue

The galaxy magnitudes, extinctions, k-corrections and stellar masses are obtained from the New York University Value Added Catalogue (NYU-VAGC; Blanton et al. 2005). The k-corrections and stellar masses are computed following the methodology of Blanton & Roweis (2007), which assume a Chabrier initial mass function. The star formation rates (SFRs) and specific star formation rates (SSFR = SFR/M∗) are from the most recent release of the spectral reductions of Brinchmann et al. (2004). Thus, to fully investigate the properties obtained from emission lines, whenever available, or determined from the 4000 Å break (D4000) when there are no emission lines or in the presence of strong contamination from active galactic nuclei (Brinchmann et al. 2004). It should be noted that SSFRs obtained from the D4000 value, which are typically values < 10−12 yr−1, are not exact measures of SSFR but should instead be taken as an upper limit. The average 2σ errors on the SFR estimates are between 0.5 and 1.0 dex, where galaxies with higher SSFRs have lower errors (Brinchmann et al. 2004).

The effects of the environment and the importance of pre-processing are probed via the quiescent fraction (hereafter fq), where fq is defined as

\[ f_q = \frac{\text{Number of galaxies with SSFR < 10}^{-11} \text{yr}^{-1}}{\text{total number of galaxies}} \] (1)

with SSFR = 10−11 yr−1 marking the division between the main sequence of star-forming galaxies and the quiescent galaxies in the SSFR-stellar mass plane (McGee et al. 2011; Wetzel et al. 2012). It should also be noted that the values in equation (1) are weighted to account for spectroscopic incompleteness using the completeness values computed in Yang et al. (2007).

1 http://www.mpa-garching.mpg.de/SDSS/DR7
An advantage of studying environmental effects via $f_{\sigma}$, rather than with mean SSFR or SSFR distributions, is that the aforementioned uncertainty in low values of SSFR derived by Brinchmann et al. (2004) do not affect the results of our analysis. To ensure that the quiescent fraction is not biased, we use a stellar mass complete sample. As a result of the magnitude limit of the SDSS survey, the stellar mass completeness limit is a function of redshift. Therefore, in order to include low-mass galaxies in our analysis, we restrict the redshift range to $z \leq 0.045$, which provides us with a sample that is complete down to $3.2 \times 10^9 M_\odot$. Analysis is performed on satellite galaxies with all ‘central’ galaxies, taken to be the most massive galaxy as identified by Yang et al. (2007), removed from our sample.

2.2 The SDSS group catalogue

Our sample consists of groups and clusters identified in SDSS by Yang et al. (2007). These authors identify groups using all galaxies in the SDSS-seventh data release (DR7) sample brighter than the survey magnitude limit of $r \leq 17.77$ and with spectroscopic completeness $>70$ per cent. The groups are identified with a halo-based group finder, which uses a traditional friends-of-friends algorithm to identify potential systems and then adds or removes members iteratively based on the mass of the dark matter halo and the assumption that the distribution of galaxies follows that of dark matter haloes, which is assumed to be a projected NFW profile (Navarro, Frenk & White 1996). The mass of the halo is determined initially from the total or characteristic luminosity ($L_{19.5}$ in Yang et al. 2007) of all the potential group members with $0.1 M_{lim} - 5 \log h^{-1} \leq -19.5$, where $0.1 M_{lim}$ is the absolute magnitude limit at the redshift of the group $k$-corrected to $z = 0.1$, and a constant mass-to-light ($M/L$) ratio of 500 $M_\odot/L_\odot$. It should be noted that only for the first iteration is a constant $M/L$ ratio used; for all subsequent iterations, the $M_{halo}/L_{19.5} - L_{19.5}$ relation from the previous iteration is used to determine the halo mass. In addition, an initial velocity dispersion and size are computed from the members of the potential group. Using this initial mass, size and dispersion, as well as an assumed NFW radial profile and a Gaussian distribution for the line-of-sight (LOS) velocities, the algorithm then adds or removes members until no further members can be added and the $M_{halo}/L_{19.5} - L_{19.5}$ relation converges.

The Yang et al. (2007) group finder identifies systems that cover a wide range of masses, from isolated galaxies to rich clusters ($M_{halo} \sim 10^{15} M_\odot$). These authors carried out performance tests of their halo-based group finder using a mock galaxy redshift survey made to mimic the SDSS-DR4 sample. The performance of the group finder was characterized by the completeness ($f_c$), defined as the number of members identified over the total number of true group members, and the contamination ($f_i$), defined as the number of interloping non-members over the total number of true members. Yang et al. (2007) found that the percentage of groups with 100 per cent completeness ranged from $\sim 93$ per cent in low-mass groups ($10^{12.5} < M_{halo} \leq 10^{13.5} M_\odot$) to 60 per cent for the most massive clusters ($10^{14.5} < M_{halo} \leq 10^{15} M_\odot$). Since the majority of systems in our sample are in the low-mass halo regime, it is expected that our groups are relatively complete. The contamination from interlopers appears to be mostly independent of halo mass. On average $\sim 65$ per cent of the systems had no contamination at all and $\sim 85$ per cent had $f_i \leq 0.5$. Interloper galaxies are typically either field galaxies or members of nearby massive groups with similar projected spatial positions but offsets along the LOS. The impact of the number of interloping galaxies needs to be explored in detail with mock catalogues from simulations, which we reserve for future work. However, we discuss possible implications of interloping galaxies for this work in Section 3.

In this analysis, we only study systems with $n_{members} \geq 10$, which is the minimum group membership for reliable substructure analysis (Hou et al. 2012). This leaves us with a total of 306 groups and 9095 member galaxies. Additionally, while our sample contains both groups (i.e. systems with $10^{12} \leq M_{halo} \leq 10^{14} M_\odot$) and clusters (i.e. systems with $M_{halo} \geq 10^{14} M_\odot$), we will refer to all systems as ‘groups’ for simplicity. In Fig. 1, we show the main properties of the groups in our sample and plot the group velocity dispersion ($\sigma_{rest}$) versus group richness ($n_{members}$) for the systems in our sample. The dispersion ($\sigma_{obs}$) is the observed velocity dispersion computed via the Gapper Estimator (Beers, Flynn & Gebhardt 1990) from all member galaxies above our stellar mass completeness limit and then corrected for redshift (i.e. $\sigma_{rest} = \sigma_{obs}/(1+z)$). The group richness is taken to be the number of group members after our stellar mass cut of $3.16 \times 10^9 M_\odot$ is applied and is therefore the number of members used in the dynamical analysis presented in this work. The majority of our sample resides in the group regime with $100 < \sigma_{rest} < 400$ km s$^{-1}$ and $10 \leq n_{members} \leq 50$.

3 IDENTIFYING SUBHALOES

In order to investigate whether pre-processed galaxies contribute to the observed morphology–/colour–density relations (e.g. Dressler 1980; Blanton et al. 2003; Balogh et al. 2004; Baldry et al. 2006; Bamford et al. 2009), we must first identify subhaloes, which we define as a collection of galaxies that occupy the same halo within a larger host group halo. One method of identifying subhaloes is to look for substructure, which is believed to be an indication of the recent accretion of galaxies or small groups of galaxies.
As in our previous work (Hou et al. 2012, 2013), we identify substructure using a modified version of the DS Test on all groups with \( N_{\text{members}} \geq 10 \) in our SDSS sample. The DS Test (Dressler & Shectman 1988) uses both spatial and LOS velocity information to identify substructure and searches for members or groups of members with kinematic properties that deviate from those of the host group. The DS \( \delta_i \)-deviation is computed for each galaxy as

\[
\delta_i = \frac{N_{\text{nn}} + 1}{\sigma^2} \left[ (v_{\text{local}} - \bar{v})^2 + (\sigma_{\text{local}} - \sigma)^2 \right],
\]  

(2)

where \( 1 \leq i \leq n_{\text{members}} \), \( N_{\text{nn}} = \sqrt{n_{\text{members}}} \) rounded down to the nearest integer in the modified version of the test (Pinkney et al. 1996; Zabludoff & Mulchaey 1998b), \( v_{\text{local}} \) and \( \sigma_{\text{local}} \) are the mean velocity and velocity dispersion of the galaxy plus its \( N_{\text{nn}} \) neighbours (as projected on the sky), and \( \bar{v} \) and \( \sigma \) are the mean velocity and velocity dispersion of the host group. Galaxies with large \( \delta_i \)-values have large kinematic deviations and could indicate new group members that have yet to adopt the kinematic properties of the host group. To determine whether a group contains significant substructure, the sum of DS deviations is computed as

\[
\Delta = \sum_{i=1}^{n} \delta_i.
\]  

(3)

Monte Carlo methods are then used to determine the probability that the computed \( \Delta \) value can be obtained from a random distribution of galaxy positions and velocities. The probability is computed by comparing the observed \( \Delta \)-value to ‘shuffled \( \Delta \)-values’, which are computed by randomly shuffling the observed velocities and then reassigning them to the observed member galaxy positions. Systems with probabilities below a given confidence level (typically 1 or 5 per cent) are identified as having significant substructure.

In our previous work, we focused on comparing groups with and without detectable substructure using the \( \Delta \) statistic (Hou et al. 2012, 2013). However, the goal of this work is to investigate the role of pre-processing, which requires the identification of individual subhaloes. A simple way to identify subhaloes involves a combined analysis of the group ‘bubble-plot’, that is a position plot of the group members where the symbols are weighted by \( \exp(\delta_i) \) and the group velocity distribution (Dressler & Shectman 1988; Dressler et al. 2013). In the bubble-plots, the size of the symbols scale with the DS \( \delta_i \)-deviation, therefore larger symbols correspond to galaxies with larger kinematic deviations from the group average. Subhalo candidates are identified in the bubble-plots as regions where several galaxies have similarly large symbols. The DS \( \delta_i \)-deviation does not take into account the sign of the galaxy velocity (equation 2); therefore, to ensure that the galaxies are also correlated in velocity-space, Dressler et al. (2013) look at the velocity distribution of the subhalo candidates. If the candidate galaxies span a small enough range in velocity (\( \lesssim 1000 \text{ km s}^{-1} \)), then Dressler et al. (2013) identify these as a subhalo.

While the bubble-plots are effective in identifying subhaloes, it is not feasible to carry out visual inspection for a large sample of groups. Therefore, we automate this process by defining subhaloes as a collection of at least three neighbouring galaxies, as projected on the plane of the sky, with \( \delta_i \geq 1.8 \) that lie within a narrow range of LOS velocities of each other. The minimum value of three neighbouring galaxies in our subhaloes corresponds to the fact that the modified version of the DS Test uses \( N_{\text{nn}} = \sqrt{n_{\text{members}}} \) (rounded down to the nearest integer) to compute \( \delta_i \) (equation 2). Since our smallest groups have \( n_{\text{members}} = 10 \) then at minimum \( N_{\text{nn}} = 3 \). The \( \delta_i \geq 1.8 \) requirement results from the observation that the average \( \delta_i \)-value is approximately 1, which can be seen in the \( \delta_i \)-distribution shown in Fig. 2. We choose the value of \( \delta_i = 1.8 \) so that \( \sim 25 \) per cent of the galaxies in our sample lie above the cut-off, but our results are not sensitive to the particular value of \( \delta_i \) that is used. The LOS velocity cut applied around each galaxy ensures that the candidate subhalo galaxies are not only close in projection on the sky, but also correlated in redshift space. Since our sample includes groups that span a wide range in halo mass and group richness (Fig. 1), we set our LOS velocity cut equal to \( \sigma_{\text{los}} \), which allows the velocity range for subhalo galaxies to scale with the mass of the host group. Theoretically, subhaloes should also span a wide range in mass; however, massive subhaloes (\( \sim 10^{13} \text{ M}_\odot \)) are likely only found in cluster-sized systems (\( \geq 10^{13} \text{ M}_\odot \); McGee et al. 2009). A constant LOS velocity cut applied to all subhaloes would either be too restrictive for rich clusters or too relaxed for lower mass groups. While we cannot reliably determine the masses of our identified subhaloes, due to the small number of galaxies within a given subhalo, visual inspection of our systems shows that subhaloes in lower mass groups typically only have a few member galaxies, while massive systems can contain subhaloes with as few as three and as many as \( \sim 10 \) member galaxies. Thus, our methodology reflects the expected range in subhalo masses for a given group halo mass.

We now compare our automated subhalo finder to the visual inspection methodology described in Dressler et al. (2013). In Fig. 3, we show bubble-plots for two example groups in our sample, Yang et al. (2007) SDSS Groups 138 (left) and 433 (right). Galaxies with \( \delta_i < 1.8 \) are indicated by black symbols, galaxies with \( \delta_i \geq 1.8 \) are indicated by blue symbols and the size of the symbols scale with \( \exp(\delta_i) \). In addition, we also indicate the galaxies identified as subhalo members using our automated algorithm (red crosses in Fig. 3). In both Groups 138 and 433, our subhalo finder clearly identifies the collection of galaxies with the largest DS \( \delta_i \)-deviations (Fig. 3). It appears that our automated method is able to identify the same subhaloes that a visual inspection would detect. This methodology was applied to all groups in our sample and we find that our algorithm systematically reproduces the subhalo population identified via visual inspection.
Figure 3. Left: declination (Dec.) versus right ascension (RA) for Yang et al. (2007) SDSS Group 138, where the symbols scale with \( \exp(\delta_i) \), often referred to as DS ‘bubble-plots’, and larger symbols correspond to larger kinematic deviations from the host group properties. Black symbols represent galaxies with \( \delta_i < 1.8 \), blue symbols represent galaxies with \( \delta_i \geq 1.8 \) and the red crosses indicate galaxies that have been identified as being part of subhalo by our automated subhalo finder. Right: same as left except for Yang et al. (2007) SDSS Group 433. The red dashed circle indicates the virial radius (i.e. \( r_{200} \)) of each system and corresponds to values of 1.27 and 1.57 Mpc for Groups 138 and 433. Subhalo galaxies identified with our automated finder (represented by red crosses) generally match those that would be identified via visual inspection (represented by blue symbols).

In Section 2.2, we mentioned that in some cases the groups identified in Yang et al. (2007) suffered from contamination from interloping galaxies. Recall that \( \sim 85 \) per cent of the systems in the Yang et al. (2007) group catalogue have anywhere between 0 and 50 per cent contamination from interloping galaxies, which were typically galaxies on the outskirts of neighbouring groups along the LOS. To determine the importance of contamination from neighbouring galaxies within our subhaloes, we computed the number of groups with subhaloes with close neighbours, defined as groups that are within \( 3r_{200} \) and \( 3\sigma \) (along the LOS) of another group. We found that only \( \sim 5 \) per cent of our sample of groups with subhaloes had close neighbours, indicating that contamination from neighbouring groups does not have a significant effect on our identification of subhalo galaxies.

4 COMPARING SUBHALO AND NON-SUBHALO GALAXIES

With our automated subhalo finder, we identify subhaloes in our group sample and find that in total \( \sim 10 \) per cent of all group galaxies reside in subhaloes. We then construct samples of subhalo and non-subhalo galaxies and in this section we compare the galaxy properties of these two populations.

4.1 Halo mass distributions

In Fig. 4, we show the differential (left) and cumulative (right) halo mass distributions for groups with no subhaloes (black solid line) and for groups with subhaloes identified with the methodology described in Section 3 (red dashed line). The halo mass distributions for groups with and without subhaloes are distinct at the \( >99 \) per cent confidence level based on the results of a two-sample KS Test. It is clear that subhaloes preferentially reside in more massive systems (Fig. 4). Almost all (\( \sim 95 \) per cent) of the groups with no identified subhaloes have halo masses \( \leq 10^{14} M_\odot \), while a significantly lower fraction (\( \sim 60 \) per cent) of groups with subhaloes lie below this halo mass. These results are generally in good agreement with results from numerical simulations and SAMs, which suggest that subhaloes are more common in more massive host groups (e.g. De Lucia et al. 2012; Bahé et al. 2013; Wetzel et al. 2013). We discuss the relationship between subhaloes and halo mass in more detail in Section 5.

4.2 Stellar mass distributions

We show the differential (left) and cumulative (right) stellar mass distributions for non-subhalo (black solid line) and subhalo (red dashed line) galaxies in Fig. 5. Although the non-subhalo and subhalo stellar mass distributions appear similar, a two-sample KS Test indicates that these two distributions likely come from distinct parent distributions at the \( >96 \) per cent confidence level. The main differences between the two populations is that the subhalo population appears to have fewer low-mass \( (\log_{10}(M_{\text{star}}/M_\odot) \lesssim 10) \) galaxies and a slightly higher fraction of more massive galaxies (Fig. 5: left). However, it should be noted that the differences in the stellar mass distributions of non-subhalo and subhalo galaxies are subtle and of the order of, at most, a few per cent (Fig. 5: right). We discuss whether these differences in the stellar mass distribution affect our results in Section 4.3.

4.3 Radial trends

In order to obtain better statistics for our analysis, we look at the stacked group properties. Our group sample consists of groups with varying sizes and so we show radial trends as a function of \( r_{\text{proj}} / r_{200} \), where \( r_{\text{proj}} \) is the projected group-centric radius and \( r_{200} \) is defined as the radius within which the average density is 200 times the
4.3.1 Radial distributions

In Fig. 6, we show the differential (left) and cumulative (right) group-centric radial distributions for galaxies in the non-subhalo (black solid line) and subhalo (red dashed line) populations in our stellar mass complete sample. From Fig. 6, we see that the two distributions differ and results from a two-sample KS Test confirm that the subhalo and non-subhalo radial distributions come from different parent distributions at the >99 percent confidence level. Additionally, we find that subhalo galaxies are preferentially found at larger radii when compared to non-subhalo galaxies (Fig. 6). The majority (~60 percent) of the galaxies in the non-subhalo population reside within the virial radius and the fraction of galaxies decreases with increasing radius. In contrast, there appears to be a dearth of subhalo galaxies close to the group core with ~60 percent of subhalo galaxies found beyond the virial radius.

critical density of the Universe. We use an approximation of the physical virial radius as defined in Carlberg et al. (1997)

$$r_{200} = \frac{\sqrt{3} \sigma_{\text{rest}}}{10H(z)}$$

(4)

where $H(z) = H_0 \sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0}}$.

Figure 4. Left: differential halo mass distribution for galaxies in groups with no identified subhaloes (black solid line) and groups with subhaloes (red dashed line). Right: same as left except we plot the cumulative halo mass distributions. It is clear that groups with subhaloes preferentially reside in more massive systems.

Figure 5. Left: normalized differential stellar mass distribution of galaxies with $M_{\text{star}} \geq 3.16 \times 10^9 M_\odot$ in the non-subhalo (black solid line) and subhalo (red dashed line) populations. Right: same as left except we plot the cumulative stellar mass distributions. The non-subhalo and subhalo stellar mass distributions appear similar; however, they are distinct at the >96 per cent confidence level based on the results of a two-sample KS Test.

4.3.1 Radial distributions

In Fig. 6, we show the differential (left) and cumulative (right) group-centric radial distributions for galaxies in the non-subhalo (black solid line) and subhalo (red dashed line) populations in our stellar mass complete sample. From Fig. 6, we see that the two distributions differ and results from a two-sample KS Test confirm that the subhalo and non-subhalo radial distributions come from different parent distributions at the >99 percent confidence level. Additionally, we find that subhalo galaxies are preferentially found at larger radii when compared to non-subhalo galaxies (Fig. 6). The majority (~60 percent) of the galaxies in the non-subhalo population reside within the virial radius and the fraction of galaxies decreases with increasing radius. In contrast, there appears to be a dearth of subhalo galaxies close to the group core with ~60 percent of subhalo galaxies found beyond the virial radius.
4.3.2 Quiescent fraction versus radius

Differences between subhalo and non-subhalo galaxies can be probed by looking at their SSFRs via the quiescent fraction ($q_\text{f}$). In Fig. 7, we show $q_\text{f}$ versus $r_{\text{proj}}/r_{200}$ for non-subhalo (black circles) and subhalo (red crosses) galaxies in our entire sample of satellite galaxies (top-left panel), for low-mass satellites ($9.5 < \log_{10}(M_{\text{star}}/M_\odot) < 10$: top-right panel), for intermediate-mass satellites ($10 < \log_{10}(M_{\text{star}}/M_\odot) < 10.5$: bottom-left panel) and high-mass satellites ($\log_{10}(M_{\text{star}}/M_\odot) > 10.5$: bottom-right panel). The group-centric radius covers a range between $0 < r_{\text{proj}}/r_{200} < 3$ and the data are plotted at the mean value of each bin, which have widths of $0.75r_{200}$. The dashed horizontal black line corresponds to the observed quiescent fraction in the field, where field galaxies are taken to be the isolated galaxies in the Yang et al. (2007) catalogue.

In the top-left panel of Fig. 7, we see that for the overall group galaxy population $q_\text{f}$ is significantly higher than in the field at all radii, indicating that group galaxies experience environmental star formation quenching out to at least three virial radii. It should be noted that the quiescent fraction for the overall satellite population is weighted to match the group galaxy stellar mass distribution for a better comparison between the field and group samples. This result is in agreement with previous observations that also find a higher quiescent fraction in groups with respect to the field as far out as $\sim 5r_{200}$ (e.g. von der Linden et al. 2010; Bahé et al. 2013; Wetzel et al. 2014). Comparing group and field galaxies at a given stellar mass, we find enhanced quenching in low- and intermediate-mass group galaxies at almost all radii. For high-mass galaxies, the non-subhalo population shows enhanced quenching at small radii ($r < 1.5r_{200}$), but have $q_\text{f}$ values similar to field on the group outskirts (Fig. 7: bottom-right panel). The high-mass subhalo galaxies show enhanced quenching, with respect to the field, closer to the group core ($r < 0.75r_{200}$) and just beyond the virial radius ($1.5 < r < 2.25r_{200}$).

As a function of radius, we see that for non-subhalo galaxies at all stellar masses the general trend is that $q_\text{f}$ decreases with increasing group-centric radius within $\sim 1.5r_{200}$ and then flattens on the group outskirts (Fig. 7). The subhalo galaxy population shows a different radial trend from non-subhalo galaxies, in both the overall satellite population and at fixed stellar mass, where $q_\text{f}$ decreases with increasing radius within $\sim 1.5r_{200}$ but then appears to increase at large radii (Fig. 7). These results indicate that on the group outskirts subhalo galaxies have experienced enhanced star formation quenching, with respect to the non-subhalo population. Within the virial radius, the quiescent fractions in non-subhalo and subhalo galaxies are similar; however, at large radii $q_\text{f}$ is higher in the subhalo population at all stellar masses, although a statistically significant difference is only observed in low- and intermediate-mass satellites beyond two virial radii (Fig. 7).

In Fig. 5, we found that the non-subhalo and subhalo stellar mass distributions differed slightly, where more massive galaxies are preferentially found in subhaloes. The quiescent fraction has also been shown to correlate with stellar mass where more massive galaxies typically have higher values of $q_\text{f}$ (e.g. Kimm et al. 2009; Wetzel et al. 2012; Hou et al. 2013; Woo et al. 2013). While this could potentially affect the overall galaxy population (Fig. 7: top-left panel), the difference in the stellar mass distributions of non-subhalo and subhalo populations is very small. More importantly, we still observe higher $q_\text{f}$ at fixed stellar in the subhalo population (Fig. 7), which suggests that environmental effects contribute to the enhanced quenching.

4.4 Separating virialized, infalling and backsplash galaxies

In Section 4.3, we showed that the $q_\text{f} \sim r_{200}$ trend for subhalo and non-subhalo galaxies differed on the group outskirts. As a result, one might naively assume that the identified subhaloes are infalling low-mass groups. However, numerous simulations have shown that backsplash galaxies can extend as far out as 2–3 virial radii (Balogh et al. 2000; Mamon et al. 2004; Gill et al. 2005; Oman et al. 2013). Additionally, results from SAMs have shown that subhaloes can survive, that is maintain the kinematic properties of the subhalo, for several orbits within the host group potential (Taylor & Babul 2004). Therefore, our goal is to distinguish between the infall and

---

**Figure 6.** Left: differential radial distributions of galaxies with $M_{\text{sat}} \geq 3.16 \times 10^9 M_\odot$ in the non-subhalo (black solid line) and subhalo (red dashed line) populations. Right: same as left except we plot the cumulative radial distributions. Subhalo galaxies are preferentially located on the group outskirts.
Pre-processing of infall galaxies

Figure 7. Top-left: quiescent fraction ($f_q$) versus $r_{proj}/r_{200}$ for all satellite non-subhalo galaxies (black circles) and all satellite subhalo galaxies (red crosses). Top-right: same as top-left except only for the low-mass ($9.5 < \log_{10}(M_{\text{star}}/M_\odot) < 10$) galaxies. Bottom-left: same as top-left except for intermediate mass ($10 < \log_{10}(M_{\text{star}}/M_\odot) < 10.5$) galaxies. Bottom-right: same as top-left except for high-mass ($\log_{10}(M_{\text{star}}/M_\odot) > 10.5$) galaxies. The data are plotted at the mean value of each radial bin, which has a width of 0.75 $r_{200}$. The dashed horizontal black line represents the quiescent fraction in isolated field galaxies in the aforementioned stellar mass bins. For the overall satellite population (top-left panel) the field quiescent fraction is weighted to match the group galaxy stellar mass distribution. Errors are computed following Cameron (2011). In general, the group galaxies (both non-subhalo and subhalo) have higher quiescent fractions than observed in the field. Additionally, on the group outskirts, $f_q$ is higher in subhalo galaxies with respect to non-subhalo galaxies, indicating that enhanced quenching has occurred.

We look to determine if our groups contain infall and backsplash satellites is to look at the distribution of the galaxy velocities ($\Delta cz$) as a function of the group velocity dispersion ($\sigma$). Backsplash galaxies will have been slowed due to dynamical friction within the group core and will therefore have low $|\Delta cz|/\sigma$ values at fixed radius (Gill et al. 2005). In contrast, infalling galaxies can have a wide range of velocities depending on their orbital parameters: though galaxies with high velocities (i.e. $|\Delta cz|/\sigma \geq 1$) are likely all infalling. Using N-body simulations, Gill et al. (2005) showed that the distribution of backsplash galaxies have a narrow centrally peaked $|\Delta cz|/\sigma$ distribution, while infalling satellites have a broader distribution with a non-zero peak. While it is difficult to separate infall and backsplash galaxies from an observed, and therefore projected, $|\Delta cz|/\sigma$ distribution, Pimbblet (2011) found that by binning the $|\Delta cz|/\sigma$ histogram into narrow bins of radius, it is possible to identify regions where infalling galaxies dominate. More specifically, these authors found that bimodality...
and/or a shift in the peak of the $|\Delta cz|/\sigma_{\text{rest}}$ distribution to larger values indicated a large infall population.

In Fig. 8, we show the $|\Delta cz|/\sigma_{\text{rest}}$ histograms in narrow radial bins for non-subhalo (left) and subhalo (right) galaxies and also list the number of galaxies in each bin. For the non-subhalo galaxies, we see that for almost all radial bins the $|\Delta cz|/\sigma_{\text{rest}}$ distribution is broad and generally centrally peaked, indicating a mixed population of virialized (for galaxies with $r < r_{200}$), infall and backsplash galaxies out to $2r_{200}$ (Gill et al. 2005). Only on the outskirts ($2 < r_{200} < 3.0$) are there signs of a large infall population, indicated by the emergence of a second peak at $|\Delta cz|/\sigma_{\text{rest}} \sim 0.5$. In contrast, the subhalo galaxies (Fig. 8: right) show signs of a strong infall population just beyond the virial radius and out to $3r_{200}$. The $|\Delta cz|/\sigma_{\text{rest}}$ distributions for galaxies between 1and$2r_{200}$ either show bimodality or an offset peak, which are both indications of a dominant infall population (Gill et al. 2005; Pimbblet 2011). There are too few subhalo galaxies in the $2.5 < r < 3r_{200}$ bin to comment on the shape of the $|\Delta cz|/\sigma_{\text{rest}}$ distribution; however, it is clear that most of the galaxies have relatively high velocities and are likely infalling.

The $|\Delta cz|/\sigma$ distributions for the subhalo galaxies shown in Fig. 8 indicate the presence of a dominant infall population; however, it is not possible to distinguish between infall and backsplash galaxies from these histograms alone. Using $N$-body simulations, both Gill et al. (2005) and Oman et al. (2013) showed that in 6-d phase-space $(x, y, z, v_x, v_y, v_z)$, the regions occupied by each population are for the most part distinct. However, once this phase-space is collapsed into observables (i.e. $x, y$ and $v_z$), projection effects tend to fill out much of the empty phase-space that separated the populations. While there is no ideal method to distinguish between infall and backsplash galaxies in observed groups, there are ways to roughly approximate regions occupied by either population. Mahajan et al. (2011) found that the fraction of virialized, infalling and backsplash galaxies occupied distinct regions in the $v_r/V_c - r/R_e$ plane, where $r$ and $v_r$ are the radial phase-space coordinates, $V_c$ is the group or cluster velocity dispersion and $R_e$ is the virial radius of the system. To distinguish between the different galaxy populations, Mahajan et al. (2011) make the following cut

$$\frac{v_r}{V_c} = -1.8 + 1.06 \left( \frac{r}{R_e} \right),$$

(5)

to separate backsplash and infall galaxies and a cut at one virial radius to separate virialized and infall galaxies. Within the virial radius there is an additional cut to separate virialized and infall galaxies, which is the mirror slope of equation (5). It should be noted that while these cuts are based on the full 6-d phase-space data, the density contours for the virialized, infall and backsplash populations occupy similar regions in projected space, though with significant overlap, and therefore contamination between the populations (Mahajan et al. 2011). While the distinct regions are not as clear in projected space, the divisions made by equation (5) allow us to approximately distinguish between infall and backsplash subhaloes, rather than assuming all subhaloes are infalling. Therefore, we apply a cut analogous to equation (5), except $v_r/V_c$ is replaced by the observable quantity $|\Delta cz|/\sigma_{\text{rest}}$ and $r/R_e$ is replaced by $r_{200}/r_{200}$. Additionally, the aforementioned classification scheme is one of five models tested by Mahajan et al. (2011). While we elect to use the best-fitting scheme, as determined by Mahajan et al. (2011), it should be noted that the fraction of backsplash galaxies can change by as much as $\sim 20$ per cent depending on the classification scheme used.

In Fig. 9, we plot $\Delta cz/\sigma_{\text{rest}}$ versus $r_{200}/r_{200}$ for our population of non-subhalo galaxies (grey crosses) and subhalo galaxies (red triangles). As in Mahajan et al. (2011), we divide the $\Delta cz/\sigma_{\text{rest}}-r_{200}$ plane into regions of virialized (Region A), infalling (Regions B) and backsplash (Region C) with equation (5) and a cut at $r_{200}/r_{200} = 1.0$. Both subhalo and non-subhalo galaxies occupy all three regions of Fig. 9, though there are some visible differences between the two populations. In particular, there are few subhalo galaxies close to the group core (also seen in Fig. 6) and there appears to be an excess of subhalo galaxies, with respect to the non-subhalo population, in the bottom-right hand corner of Fig. 9. This area corresponds to the region occupied by only infalling galaxies in the full 6-D phase-space diagram shown in Mahajan et al. (2011).
5 HOW SIGNIFICANT IS PRE-PROCESSING?

To determine the importance of pre-processing in groups and clusters, we first quantify the fraction of galaxies that reside in subhaloes ($f_{\text{sub}}$) defined as the number of galaxies in identified subhaloes over the total number of group members. In Fig. 11 (left), we plot $f_{\text{sub}}$ versus host group halo mass ($M_{\text{halo}}$), where $M_{\text{halo}}$ is the luminosity-based halo mass from Yang et al. (2007). The data points correspond to values of $f_{\text{sub}}$ computed for individual systems and the horizontal red lines represent the mean value of $f_{\text{sub}}$ computed for each halo mass bin, which has a width of 0.5 dex. It can be clearly seen in Fig. 11 (left) that for halo masses below $\sim 10^{13.2} M_{\odot}$ our automated subhalo finder does not identify any subhaloes, which can also be seen in Fig. 4. Between $10^{13.2} \lesssim M_{\text{halo}} \lesssim 10^{14.2} M_{\odot}$, there
is a mixture of systems with and without subhaloes; while the most massive clusters \((M_{\text{halo}} \gtrsim 10^{14.2} M_\odot)\) all contain subhaloes (Figs 4 and 11: left). Although there is significant scatter in the \(f_{\text{sub}}\) values of individual systems, the mean values (red lines in Fig. 11: left) appear to show a trend with halo mass, where more massive haloes have a higher fraction of subhaloes. However, more data is required to confirm the observed correlation between \(f_{\text{sub}}\) and halo mass. A similar dependence of \(f_{\text{sub}}\) on halo mass is also seen in both SAMs (De Lucia et al. 2012) and hydrodynamical simulations (Bahé et al. 2013).

The results of Fig. 11 (left) provide information about the relationship between \(f_{\text{sub}}\) and halo mass; however, as shown in Section 4.4, the subhalo population contains a mix of both infalling and backsplash galaxies. In order to better estimate the importance of pre-processing, we must examine the fraction of infalling subhaloes \((f_{\text{sub, infall}})\) defined as the number of infalling subhalo galaxies over the total number of infalling galaxies, where the galaxies are classified with the divisions shown in Fig. 9. A value of \(f_{\text{sub, infall}} = 0\) implies that all of the infalling galaxies are either accreting directly from the field or are part of a subhalo that is not identified by our algorithm (i.e. small or not very kinematically distinct subhaloes). In addition, \(f_{\text{sub, infall}} = 0\) could also indicate that either all of the subhaloes in that group are in the backsplash population or that the group does not contain any identified subhaloes. A value of \(f_{\text{sub, infall}} = 1\) indicates that all of the infalling galaxies are subhaloes.

In Fig. 11 (right), we plot \(f_{\text{sub, infall}}\) versus \(\log_{10}(M_{\text{halo}}/M_\odot)\), where the data points indicate individual systems and the red lines define the mean value within each halo mass bin. There are many groups (~85 per cent) with \(f_{\text{sub, infall}} = 0\). However, similar to the fraction of groups with subhaloes (Fig. 11: left) the number of groups with infalling subhaloes shows a trend with halo mass, where \(f_{\text{sub, infall}}\) increases with increasing halo mass. For groups \((M_{\text{halo}} < 10^{14} M_\odot)\), the sample is dominated by systems with \(f_{\text{sub, infall}} = 0\), which results in mean \(f_{\text{sub, infall}}\) values ~5 per cent. For clusters with \(10^{14} < M_{\text{halo}} < 10^{15} M_\odot\), the mean value of \(f_{\text{sub, infall}}\) is ~10 per cent, which is in good agreement with the fraction of pre-processed galaxies predicted by Berrier et al. (2009) for a similar halo mass range. Only for the most massive clusters in our sample \((M_{\text{halo}} > 10^{15} M_\odot)\) do we find that a significant fraction (~25 per cent) of the infall population reside in subhaloes. Taking these average cluster values, we find that our results are somewhat lower than the values predicted by the SAMs of McGee et al. (2009) and De Lucia et al. (2012), who found that the fraction of galaxies that accrete on to clusters (with \(\log_{10}(M_{\text{halo}}/M_\odot) \gtrsim 14\)) as members of groups with \(\log_{10}(M_{\text{halo}}/M_\odot) \gtrsim 10^{13}\) ranges between ~25 and 45 per cent. A possible explanation for the discrepancy between our observed value and values predicted by some SAMs (McGee et al. 2009; De Lucia et al. 2012) is that our automated subhalo finder cannot detect smaller and/or less kinematically distinct subhaloes, and thus our fraction of subhaloes is likely a lower limit (see Hou et al. 2012, for a discussion on the limitations of the DS Test). Also, our sample of cluster-sized systems is small (~14 per cent) and since there is significant scatter on the individual values of \(f_{\text{sub, infall}}\) (Fig. 11: right), it is possible that our computed mean value may underestimate the true fraction of infalling subhaloes in clusters. Additionally, we note that our automated subhalo finder can only detect galaxies that are currently in subhaloes; however, some galaxies may have been in a subhalo in the past and may have also been pre-processed. Our methodology would miss such systems, and therefore our computed fraction of infalling galaxies in subhaloes would again underestimate the true value.

Based on the results shown in Figs 7 and 10, we conclude that the enhanced quenching in subhaloes observed on the group outskirts \((\gtrsim 2r_{200})\) is mainly a result of the pre-processing of infalling subhalo galaxies. While we do observe a slightly enhanced quiescent fraction in the non-subhalo population with respect to the field at large radii (Fig. 7), this can potentially be explained by...
the ~20% ejected satellite within the range $2 \lesssim r_{200} < 3.0$ (see also Wetzel et al. 2014). The main result of Fig. 7 is that we observe additional quenching in the subhalo population, with respect to both the field and non-subhalo populations, which hints at a quenching mechanism related to the subhalo itself (e.g. pre-processing). Our results also indicate the importance of pre-processing to be a function of host halo mass. For group-sized systems, pre-processing does not play a significant role in star formation quenching; however, for the cluster-sized systems, and in particular clusters with $M_{\text{halo}} > 10^{14.5} M_\odot$, a significant fraction of the member galaxies appear to have had their star formation quenched in smaller haloes prior to accretion on the final (observed) system. Unfortunately, it is not possible to further divide the results shown in Fig. 7 by halo mass as the uncertainties become too large to draw any meaningful conclusions and more data, especially massive cluster data, are needed. However, in a similar analysis of rich clusters, Dressler et al. (2013) found that the fraction of passive and post-starburst galaxies was significantly higher in their identified infalling groups and they also concluded that ‘substantial’ pre-processing had occurred. The importance of pre-processing has also been studied using hydrodynamical simulations by Bahé et al. (2013). These authors found a similar relationship between pre-processing and halo mass where more massive systems had a much higher fraction of galaxies that had been pre-processed.

6 CONCLUSIONS

We have looked at the infall and backsplash subhalo populations in SDSS-DR7 groups and clusters, using a sample of satellite galaxies, which is complete to $M_{\text{sat}} = 3.16 \times 10^9 M_\odot$. The aim of this work is to investigate the importance of pre-processing in group and cluster galaxies. The DS Test was used to identify subhalo galaxies and we followed the methodology of Mahajan et al. (2011) to classify virialized, infall and backsplash galaxies. The main results of this analysis are as follows.

(i) Subhaloes preferentially reside in massive systems and at large group-centric radii.

(ii) The stellar mass distributions of non-subhalo and subhalo galaxies are marginally distinct, where subhaloes have, on average, slightly more massive galaxies.

(iii) Low- and intermediate-mass group galaxies out to $3 r_{200}$ and high-mass satellites close to the group core show enhanced SF quenching with respect to the field.

(iv) On the group and cluster outskirts, within the range $2 \lesssim r_{200} < 3.0$, $f_q$ is higher in galaxies that reside in subhaloes than for the overall satellite galaxy population at all stellar masses.

(v) As a function of radius, the percentages of infall and backsplash galaxies do not differ between non-subhalo and subhalo galaxies.

(vi) Below halo masses of $\sim 10^{13.2} M_\odot$, all groups do not contain any detected subhaloes, while more massive haloes show a scatter in the values in both the fraction of subhalo galaxies ($f_{\text{sub}}$) and the fraction of infalling galaxies in subhaloes ($f_{\text{sub, infall}}$). Additionally, there appears to be a trend between $f_{\text{sub}}$ and $f_{\text{sub, infall}}$ with halo mass, where more massive haloes have both more subhaloes and more infalling subhaloes.

The observed enhanced quenching in infalling subhalo galaxies suggests that pre-processing does play a role in galaxy evolution; however, the significance of pre-processing depends on halo mass. Pre-processing does not appear to be the dominant mechanism in groups and low-mass clusters ($M_{\text{halo}} \lesssim 10^{14.5} M_\odot$), but it does play a significant role in producing the observed quenched fraction in massive clusters with $M_{\text{halo}} > 10^{14.5} M_\odot$.

ACKNOWLEDGEMENTS

We would like to thank the anonymous referee for their many helpful comments and suggestions. AH, LCP and WEH would like to thank the National Science and Engineering Research Council of Canada (NSERC) for funding. We would like to thank X. Yang for making their SDSS-DR7 group catalogue publicly available, the NYU-VAGC team for publication of their SDSS catalogue and J. Brinchmann for publication of their SDSS SFRs. This work would not have been possible without these public catalogues.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society and the Higher Education Funding Council for England. The SDSS web site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

REFERENCES

Bahé Y. M., McCarthy I. G., Balogh M. L., Font A. S., 2013, MNRAS, 430, 3017

Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469

Balogh M. L., McGee S. L., 2010, MNRAS, 402, L59

Balogh M. L., Navarro J. F., Morris S. L., 2000, ApJ, 540, 113

Balogh M. L., Baldry I. K., Nichol R., Miller C., Bower R., Glazebrook K., 2004, ApJ, 615, L101

Bamford S. F. et al., 2009, MNRAS, 393, 1324

Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32

Berrier J. C., Stewart K. R., Bullock J. S., Purcell C. W., Barton E. J., Wechsler R. H., 2009, ApJ, 690, 1292

Blanton M. R., Roweis S., 2007, AJ, 133, 734

Blanton M. R. et al., 2003, ApJ, 594, 186

Blanton M. R. et al., 2005, AJ, 129, 2562

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Cameron E., 2011, PASA, 28, 128

Carlberg R. G. et al., 1997, ApJ, 485, L13

Carollo C. M. et al., 2013, ApJ, 776, 71

De Lucia G., Weinmann S., Poggianti B. M., Aragón-Salamanca A., Zaritsky D., 2012, MNRAS, 423, 1277

Dressler A., 1980, ApJ, 236, 351

Dressler A., Shectman S. A., 1988, AJ, 95, 985

Dressler A., Oemler A., Jr, Poggianti B. M., Gladders M. D., Abraham L., Vulcani B., 2013, ApJ, 770, 62

George M. R. et al., 2011, ApJ, 742, 125
