Probing the galaxy–halo connection with total satellite luminosity

Jeremy L. Tinker [1]★ Junzhi Cao,[1] Mehmet Alpaslan [1] Joseph DeRose,[2,3,4,5,6] Yao-Yuan Mao [7,8,9] and Risa H. Wechsler[2,3]

[1]*Center for Cosmology and Particle Physics, Department of Physics, New York University, New York, NY 10003, USA*
[2]*Kavli Institute for Particle Astrophysics and Cosmology and Department of Physics, Stanford University, Stanford, CA 94305, USA*
[3]*Department of Particle Physics and Astrophysics, SLAC National Accelerator Laboratory, Stanford, CA 94305, USA*
[4]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
[5]*Berkeley Center for Cosmological Physics, Department of Physics, University of California, Berkeley, CA 94720, USA*
[6]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 93720, USA*
[7]*Department of Physics and Astronomy Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA*
[8]*Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA*
[9]*Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, Pittsburgh, PA 15260, USA*

ABSTRACT

We demonstrate how the total luminosity in satellite galaxies is a powerful probe of dark matter haloes around central galaxies. The method cross-correlates central galaxies in spectroscopic galaxy samples with fainter galaxies detected in photometric surveys. Using models, we show that the total galaxy luminosity, $L_{\rm sat}$, scales linearly with host halo mass, making $L_{\rm sat}$ an excellent proxy for $M_{\rm h}$. $L_{\rm sat}$ is also sensitive to the formation time of the halo. We demonstrate that probes of galaxy large-scale environment can break this degeneracy. Although this is an indirect probe of the halo, it yields a high signal-to-noise ratio measurement for galaxies expected to occupy haloes at $<10^{12}$ $M_\odot$, where other methods suffer from larger errors. In this paper, we focus on observational and theoretical systematics in the $L_{\rm sat}$ method. We test the robustness of our method of finding central galaxies and our methods of estimating the number of background galaxies. We implement this method on galaxies in the
Sloan Digital Sky Survey (SDSS) data, with satellites identified in fainter imaging data. We find excellent agreement between our theoretical predictions and the observational measurements. Finally, we compare our $L_{\rm sat}$ measurements to weak lensing estimates of $M_{\rm h}$ for red and blue subsamples. In the stellar mass range where the measurements overlap, we find consistent results, where red galaxies live in larger haloes. However, the $L_{\rm sat}$ approach allows us to probe significantly lower mass galaxies. At these masses, the $L_{\rm sat}$ values are equivalent. This example shows the potential of $L_{\rm sat}$ as a probe of dark haloes.

1 INTRODUCTION

The connection between galaxies and dark matter haloes is critical both for our understanding of galaxy formation and for constraining cosmology (see Wechsler & Tinker 2018 for a recent review). There are myriad approaches to observationally constraining the galaxy–halo connection. These approaches generally separate into two categories: indirectly inferring the connection statistically, and directly probing the dark matter halo masses through their gravitational potential. In this paper, we use a method that is complementary to, but distinct from, these direct approaches. This method, which utilizes the total luminosity in satellite galaxies, $L_{\rm sat}$, is a direct observable, but it is an indirect probe of halo mass. However, it yields a higher signal-to-noise ratio measurement than other methods and is robust to significantly lower mass dark matter haloes. Using the luminosity in satellite galaxies only, as opposed to including the light from the central galaxy, makes the approach far more sensitive to the properties of haloes around galaxies at the Milky Way mass scale and below. This technique has been used to probe haloes around isolated galaxies. In this paper, we enhance this method by incorporating new photometric data, new tests of both observational and theoretical systematics, and a new method of central galaxy selection. We will compare our results to previous studies.

Indirect, statistical approaches to quantifying the galaxy–halo connection usually focus on galaxy clustering and number densities (e.g. Zehavi et al. 2005, 2011; Tinker et al. 2007; Zheng, Coil & Zehavi 2007; Rodríguez-Puebla et al. 2015) or from galaxy group catalogues (Yang, Mo & van den Bosch 2008; Reddick et al. 2013; Sinha et al. 2018). These analyses parametrize the relationship between galaxies and haloes with a halo occupation distribution (HOD) and then constrain the free parameters from the data.[1] However, these

★ E-mail: jlt12@nyu.edu

[1]One can also use hybrid approaches that incorporate satellite kinematic data or weak lensing measurements with clustering and abundances to constrain the free parameters of the halo occupation models (More et al. 2011; Leauthaud et al. 2012; Tinker et al. 2013; Zu & Mandelbaum 2015; Lange et al. 2019). But these hybrid approaches are still distinct from direct approaches that only use observables sensitive to the dark matter potential,

indirect methods are most effective for halo mass scales above $M_{\rm h} \approx 10^{12}$ M$_\odot$, or for galaxies above $M_* \approx 10^{10.5}$ M$_\odot$. In these clustering-based approaches, halo clustering becomes independent of mass at $M_{\rm h} < 10^{12}$ M$_\odot$, thus the observed clustering amplitude carries little information about the host halo masses of central galaxies below these mass scales. This lack of constraining power is evident when comparing different analyses of low-redshift Sloan Digital Sky Survey (SDSS) galaxy samples to determine the galaxy–halo connection for red and blue galaxy subsamples. Wechsler & Tinker (2018) compiled recent results of the relative stellar masses of red and blue central galaxies at fixed halo mass (cf. their fig. 8, which presents $M_{*\rm red}/M_{*\rm blue}$ as a function of $M_{\rm h}$). Below 10^{12} M$_\odot$, there is nearly an order of magnitude difference in $M_{*\rm red}/M_{*\rm blue}$, even though most of the analyses utilize the same SDSS data set. Because of the lack of constraining power in the data, the results reflect the assumptions made in the modelling rather than physical reality.

Direct probes of the gravitational potential of dark haloes are through gravitational lensing and satellite kinematics (e.g. Conroy et al. 2005; Norberg, Frenk & Cole 2008; Hudson et al. 2015; Mandelbaum et al. 2016). For direct probes, when applied to the large spectroscopic surveys with the SDSS, the signal-to-noise ratio of direct probes degrades rapidly at $M_{\rm h} < 10^{12}$ M$_\odot$. Deeper and higher resolution imaging makes measurements of lensing masses tenable at lower masses, but at the cost of sample size.

Thus, there is a need for a complementary method of probing dark matter haloes around lower mass galaxies; one that is a direct observable, uninfluenced by model assumptions of halo occupation methods but has the statistical precision to probe the haloes around low-mass galaxies. The halo mass scale of 10^{12} M$_\odot$ is auspicious for studies of galaxy formation. The transition from gas accretion being predominantly cold to exclusively hot occurs at or below this scale (Kereš et al. 2005, 2009; Dekel & Birnboim 2006). From abundance matching analyses, this is also the scale at which most star formation occurs throughout the history of the Universe (Behroozi, Wechsler & Conroy 2013b). As a result, all approaches to constrain the galaxy halo connection find that that this mass scale is the *pivot point* of the stellar-to-halo mass relation (SHMR), where $M_*/M_{\rm h}$ is maximal (see Wechsler & Tinker 2018 and references therein). Thus, being able to directly probe the relationship between galaxies and haloes at scales of $M_{\rm h} \sim 10^{11}$–10^{12} M$_\odot$ would open a door to our understanding of galaxy formation that has to this point remained closed.

The $L_{\rm sat}$ method probes dark matter haloes by measuring the total amount of light in satellite galaxies within a dark halo. All galactic dark matter haloes, regardless of their mass, contain significant amounts of substructure within them. These substructures, which we refer to as subhaloes, also can contain galaxies. Simulations show that the number of galaxy-occupied subhaloes should scale roughly linearly with host halo mass (e.g. Kravtsov et al. 2004; Reddick et al. 2013). For a spectroscopic survey like SDSS, the majority of satellites in Milky Way-type haloes are significantly below the magnitude limit to be selected for spectroscopy. Recently, the Satellites Around Galactic Analogs (SAGA) Survey (Geha et al. 2017) performed a detailed study of eight Milky Way-like galaxies, searching for satellite galaxies at significantly fainter magnitudes. Each galaxy contained faint satellites, with the number ranging from 2 to 10 objects. We use a complementary approach here: rather than assign specific satellites to specific systems, we stack sets of spectroscopic galaxies to measure the mean number of faint satellites within the stacked sample, detected in deeper imaging data.

More than constraining the mean relationship between M_* and $M_{\rm h}$, the $L_{\rm sat}$ method we present here can uncover secondary trends in the galaxy–halo connection. Because $L_{\rm sat}$ is not a direct observable of the gravitational potential, an independent calibration would be required to convert $L_{\rm sat}$ to $M_{\rm h}$. However, comparing the relative values of $L_{\rm sat}$ of galaxies is a robust observable, and presents a unique test for galaxy formation models. For example, as presented in Wechsler & Tinker (2018), it is an open question whether the scatter in $M_{\rm h}$ at fixed M_* correlates with any other galaxy or halo properties, such as galaxy size, stellar velocity dispersion, or morphology. In a companion paper (Alpaslan & Tinker 2020), we present first results quantifying these correlations using the $L_{\rm sat}$ method.

An important caveat with this method is that, when measuring relative values of $L_{\rm sat}$, one is not necessary measuring relative values of $M_{\rm h}$. It is well know that the amount of substructure in a halo is correlated with the formation history of that halo, an effect that is part of 'halo assembly bias' (e.g. Wechsler et al. 2006; Gao & White 2007; Mao, Williamson & Wechsler 2015). This makes $L_{\rm sat}$ a *more* insightful probe of the galaxy–halo connection than simply probing dark matter mass itself. Ever since the first discoveries of halo assembly bias, the persistent and controversial question has been whether this effect propagates into the galaxy formation process (see the discussion in Wechsler & Tinker 2018). If secondary properties of central galaxies correlate with halo formation history, these properties will correlate with large-scale environment in a distinct manner. The $L_{\rm sat}$ observable, combined with these spatial tests (e.g. Abbas & Sheth 2006; Tinker et al. 2008, 2017, 2018b; Peng et al. 2010; Zu & Mandelbaum 2018; Walsh & Tinker 2019; Wang et al. 2019), offers to ability not only to tell if central galaxies of the same mass, but different secondary properties, live in different haloes, but can discriminate between differences in halo mass or in halo formation history.

In this paper, we present the following:

(i) theoretical predictions for $L_{\rm sat}$;
(ii) tests of our observational methods;
(iii) first results of $L_{\rm sat}$ versus $M_{\rm h}$ and $L_{\rm sat}$ versus M_*.

The method is based on stacking objects of a common set of properties. The objects the stacks are centred on – spectroscopic galaxies with a likelihood of being central galaxies within their host haloes – we will refer to as 'primary' objects. Once the stack is collated, we measure the luminosity function of faint objects around the primaries, assuming all galaxies are at the redshift of the primary. We subtract off the estimated background contribution to this luminosity function from objects projected along the line of sight. Whatever remains after background subtraction are satellite galaxies associated with the haloes of the primary objects. This technique has been used to quantify satellite populations in massive objects, such as galaxy clusters (e.g. Hansen et al. 2009) and luminous red galaxies (Tal et al. 2012). With these types of objects, there is minimal contamination by misclassification of satellite galaxies to be in the primary galaxy sample. Wang & White (2012) and Sales et al. (2013) went to fainter primary objects, circumventing the problem of satellite contamination by restricting the analysis to isolated primary galaxies. Lan, Ménard & Mo (2016) used a galaxy group finder to identify which galaxies are central galaxies, as well as to assign halo masses to the groups themselves. Using SDSS imaging data, they quantified the conditional luminosity function (CLF) of faint satellite galaxies as a function of halo mass. In this work, we extend this method to much fainter primary objects than Wang & White (2012)

without any constraints based on halo occupation models and the number density of galaxies.

Table 1. Volume-limited SDSS samples. The first column sets the magnitude threshold at the maximum redshift, while the stellar mass listed is the limit of 98 per cent completeness. Galaxies must be brighter and more massive than the limits listed to be in the sample. Column 4 lists the minimum halo mass returned by the group finder.

$M_r - 5\log h$	$z_{\rm max}$	$\log M_*/(\mathrm{M}_\odot)$	Min. $M_{\rm h}$ ($\mathrm{M}_\odot$)	$N_{\rm gal}$
−17.48	0.033	9.7	2.4×10^{11}	16 975
−18.32	0.047	10.1	3.3×10^{11}	30 144
−19.04	0.065	10.5	8.1×10^{11}	41 659

using deeper imaging data and more efficient methods of identifying central galaxies. In contrast to the Lan et al. (2016) study, the primary goal of this work is to use faint satellites to infer properties of the host dark matter haloes.

Our theoretical predictions are constructed using the abundance matching framework (see e.g. Wechsler & Tinker 2018), coupled with high-resolutions N-body simulations that resolve substructure even within low-mass haloes. The tests of the observational methods focus on two aspects: defining a sample of central galaxies within the spectroscopic galaxy sample, and estimating the density of background imaging galaxies. For measurements on survey data, primary objects are selected from the main galaxy sample (MGS) of the SDSS (Strauss et al. 2002). Secondary objects come from the Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Surveys (DLIS; Dey et al. 2019), which have an r-band depth of $r \sim 24$, more than 6 mag fainter than the SDSS spectroscopic sample.

The outline of this paper is as follows. In Section 2, we detail the observational data, both imaging and spectroscopic, that will be utilized to make $L_{\rm sat}$ measurements. In Section 3, we will construct theoretical models for $L_{\rm sat}$ from the abundance matching framework, focusing not just on how $L_{\rm sat}$ scales with $M_{\rm h}$ but also how other halo properties correlate with satellite luminosity. In Section 4, we will evaluate the efficacy of our methods for finding central galaxies in spectroscopic samples, but volume- and flux-limited. In Section 5, we present details of applying the method to observational data, presenting tests of the method on mock galaxy distributions. In Section 6, we present our first results on the $L_{\rm sat}$ of dark matter haloes around SDSS central galaxies. In Section 7, we summarize and discuss the results.

2 DATA

Our analysis combines spectroscopic data from SDSS and imaging data from the DLIS. First, we will describe how we construct samples of central galaxies from SDSS data. Central galaxies come from both volume-limited catalogues and the full flux-limited Sloan catalogue. Second, we describe the imaging data, including the cuts employed to create the sample of secondary galaxies.

2.1 SDSS central galaxies

We use the spectroscopic data from the SDSS MGS to find central galaxies (Strauss et al. 2002). Specifically we use the Data Release 7.2 (DR7.2) of the New York University Value-Added Galaxy Catalog (NYU-VAGC; Blanton et al. 2005a).

We use two complementary methods to construct samples of central galaxies: volume-limited group catalogues, and a flux-limited catalogue of central galaxies. For the former, we use a combination of three different volume-limited samples, listed in Table 1. This table also shows the minimum halo mass in the sample, as determined by our galaxy group catalogue. The group-finding algorithm we use is detailed in Tinker, Wetzel & Conroy (2011), based on the halo-based group finder of Yang et al. (2005), and further vetted in Campbell et al. (2015). The standard implementation of the group finder yields central galaxy samples with a purity and completeness of ∼85–90 per cent (Tinker et al. 2011), but we will discuss methods of using the group finder information to construct higher purity samples with limited loss of completeness.

We also apply a central-galaxy finding algorithm to the full flux-limited SDSS MGS. The algorithm is described in detail in Appendix A. In short, the approach relies of pre-tabulated SHMRs to assign haloes to observed galaxies, then uses the same probabilistic approach as that used in the group catalogue to determine whether a given galaxy is a satellite within a larger halo. This method, although approximate, yields central galaxy samples consistent with that of the volume-limited group catalogue. But the use of pre-tabulated SHMRs does not require that the sample be volume limited. Additionally, it does not require the sample to be statistically representative, which is a requirement of the group catalogue. This allows the central finder to yield robust results at high redshifts, where only the brightest galaxies exist, and at low stellar masses, which are only found in small numbers at the lowest redshifts.

In this paper, we will use different projected apertures within which to measure $L_{\rm sat}$. We define $L_{\rm sat}^{R}$ as the satellite luminosity within the projected virial radius of the halo, $R_{\rm vir}$. We define $L_{\rm sat}^{50}$ and $L_{\rm sat}^{100}$ as satellite luminosity within fixed comoving projected apertures of 50 and 100 h^{-1} kpc, respectively. To measure $L_{\rm sat}^{R}$ from SDSS data, we use the volume-limited group catalogues because the halo mass estimates are reasonably accurate. To measure $L_{\rm sat}^{50}$ as a function of M_*, we use the flux-limited central catalogue because it has better statistics at both low and high stellar masses.

2.2 Legacy Survey imaging data

The DESI Legacy Imaging Surveys (DLIS) is a combination of three different imaging surveys. At declinations below $+32^\circ$, in both the North Galactic Cap (NGC) and South Galactic Cap (SGC), data come from the Dark Energy Camera (DECam) instrument on the 4-m Blanco Telescope (Flaugher et al. 2015). This includes g, r, and z-band data, referenced as the DECam Legacy Survey (DECaLS). In the SGC above $+32^\circ$, data come from the Beijing–Arizona Sky Survey (BASS) on the Bok Telescope (Zhou et al. 2018). This includes g- and r-band imaging. The z-band imaging over the same area of the sky comes from the Mayall z-band Legacy Survey (MzLS; Zhou et al. 2018). Once completed, the total area covered in the DLIS will be 14 000 deg^2, with DECaLS comprising 9000 deg^2 and the combined northern facilities supplying the additional 5000 deg^2 of coverage to complete the footprint. The DLIS footprint is largely coincident with the footprint of the final SDSS imaging footprint.

For this paper we use Data Releases 6 and 7 (DR6 and DR7). DR6 comprises the NGC surveys, BASS and MzLS, while DR7 is the latest DECaLS release. DR6 covers 3823 deg^2 with at least one pass in each imaging band, and 1441 deg^2 of three-pass coverage in all bands, which is the full depth of the survey. DR7 covers nearly 9300 deg^2 with at least one pass in each imaging band, and 4355 deg^2 of three-pass coverage in all bands. Although our fiducial results use only r-band data to measure $L_{\rm sat}$, we require at least one pass in all three bands for data to be part of our analysis. Single-pass depth of the survey is ∼23.3 in r band, with a full three-pass depth of ∼23.9, although the exact depth of the survey fluctuates across the footprint at fixed pass number. We will detail to quality cuts imposed on the data in Section 5.

Table 2. *N*-body simulations.

Name	Ω_m	Ω_b	σ_8	H_0	L_{box}	N_p	m_p
C250	0.295	0.047	0.834	68.8	250	2560^3	7.63×10^7
C125	0.286	0.047	0.82	70.0	125	2048^3	1.8×10^7

Before doing any measurements, we first perform quality cuts on the data. Starting from the DLIS sweep files, we perform the following cuts on the data:

remove all objects with $type == $ 'PSF';
remove all objects with $nobjs == 0$ in the g, r, or z bands;
remove all objects with $flux\_ivar <= 0$ in the g, r, or z bands;
remove all objects with $fracmasked > 0.6$ in the g, r, or z bands.

The `fracmasked` keyword specifies the total fraction of pixels in an object that are masked out. Masking can be due to bright stars, saturated pixels, and a number of other minor occurrences that are detailed in the DLIS documentation. To account for the minor differences in the geometry of the DLIS survey and the footprint of the SDSS MGS, we only include SDSS central galaxies that lie far enough within the DLIS footprint such that an annulus with $R = 3R_{vir}$ of the estimated dark matter halo is completely within the survey area covered by at least one observation in all of g, z, and r bands.

3 THEORETICAL PREDICTIONS FOR L_{SAT}

Before testing the methodology of measuring L_{sat} on mocks and data, we first need a baseline expectation for L_{sat} as a quantity that correlates with dark matter haloes. In this section, we present our framework for constructing these theoretical models, and test not just how L_{sat} scales with M_h, but also how L_{sat} scales with M_* and how it correlates with secondary halo properties other than mass.

3.1 Numerical simulations and methods

To make theoretical predictions for L_{sat} in the context of Λ cold dark matter (ΛCDM) structure formation, we combine high-resolution *N*-body simulations with abundance matching models. Table 2 shows the properties of the two simulations used to make predictions here. Both use the ROCKSTAR code (Behroozi, Wechsler & Wu 2013a) to identify haloes and CONSISTENT TREES (Behroozi et al. 2013c) to track the merger histories of individual galaxies.

The smaller volume simulation, C125, has ∼4 times better mass resolution than the larger volume C250. We require that each simulation has enough resolution to track the subhaloes that would contain satellites down to an absolute magnitude of $M_r - 5\log h = -14$, which is the current limits of our observational results. Comparison between the two simulations at fixed host halo mass shows a small offset in the total L_{sat} values at $M_h \lesssim 10^{12}\,M_\odot$ of roughly 0.1–0.2 dex. We will use C125 to show predictions of the mean trends of L_{sat}, but use C250 to predict clustering results because the larger volume is necessary to the reduce noise in clustering results.

To connect galaxy luminosity and stellar mass to dark matter haloes, we use abundance matching (see e.g. Wechsler & Tinker 2018). To assign M_r to haloes and subhaloes, we use the Blanton et al. (2005b) luminosity function, which specifically corrects for incompleteness in SDSS observations of low-luminosity, low surface brightness galaxies. We connect M_r to M_{peak}, the peak halo mass throughout a (sub)halo's history assuming a scatter of 0.2 dex in luminosity at fixed M_{peak}.[2] To assign M_* to haloes and subhaloes, we use the stellar mass function presented in Cao et al. (2020), which utilized the principal component analysis (PCA) galaxy stellar masses of Chen et al. (2012). The stellar mass function for SDSS DR7 is also presented in Cao et al. Although we require that our abundance matching in luminosity is robust down to very low luminosities, for stellar mass we only require that the abundance matching is accurate down to $M_* \sim 10^9\,M_\odot$, which is the lower limit of our sample of central galaxies in SDSS.

3.2 Scaling of L_{sat} with M_h and M_*

Fig. 1 shows the predictions for L_{sat} from the C125 simulation, binned both as a function of M_h and M_*. Results are for all satellites brighter than $M_r - 5\log h < -14$. In both panels, the results are shown for our three different apertures.

The left-hand side of Fig. 1 shows the results as a function of M_h. For $L_{sat}(R_{vir})$, the total satellite luminosity scales roughly as a power law, with a scaling slightly steeper than linear. This is expected, as subhalo populations are mostly self-similar when one scale up the host halo mass (Gao et al. 2004), and the number of subhaloes at fixed M_{peak} scales linearly with host halo mass (Kravtsov et al. 2004). Between $M_h = 10^{11}$ and $10^{12}\,M_\odot$, the luminosity of the central galaxy increases by a factor of ∼4, while L_{sat} increases by a factor of ∼20, making is a much more sensitive diagnostic of halo mass than either the luminosity of the central galaxy or the combined luminosity of both central and satellite galaxies together.

For the 50 and 100 h^{-1} kpc apertures – L_{sat}^{50} and L_{sat}^{100}, respectively – the results at $M_h < 10^{12}\,M_\odot$ are mostly unchanged, but at higher masses the trends of L_{sat} flatten out due to the larger cross-sections of these haloes.

The right-hand side of Fig. 1 shows the same results, but now binning by M_* rather than by M_h. As expected by the scaling between M_* and M_h, the trend of L_{sat} with M_* is not as steep as before, but there is still a clear power-law dependence of $L_{sat} \sim {M_*}^{0.6}$ for $M_* < 10^{10.5}\,M_\odot$. At larger masses, the trend becomes much steeper as the SHMR flattens out, meaning that a small change in galaxy mass corresponds to a larger change in M_h. We note that using fixed apertures does not change the scaling of L_{sat} with M_* for $M_* < 10^{10.5}\,M_\odot$.

3.3 Dependence of L_{sat} on secondary halo properties

Haloes exhibit assembly bias (e.g. Wechsler et al. 2006; Gao & White 2007, and see Mao, Zentner & Wechsler 2018; Salcedo et al. 2018; Mansfield & Kravtsov 2020 for more recent treatments). At fixed mass, certain secondary properties of a halo are correlated with its large-scale environment, and thus impact halo clustering. One known property that exhibits halo assembly bias is the amount of substructure within the halo (Zentner et al. 2005; Wechsler et al. 2006; Gao & White 2007; Mao et al. 2015). Early-forming haloes have lower numbers of subhaloes, N_{sub}, because accreted subhaloes have had a longer amount of time to be tidally disrupted or merge with the host halo due to dynamical friction. Late-forming haloes have had more recent accretion events, thus they will have a surplus of substructure. Halo formation history correlates strongly with large-scale environment, such that early-forming low-mass haloes are in

[2] We note that there is marginal (<0.1 dex) difference in the results when using V_{peak} – the largest value of the halo's maximum circular velocity during its evolution – as the halo parameter to abundance match to.

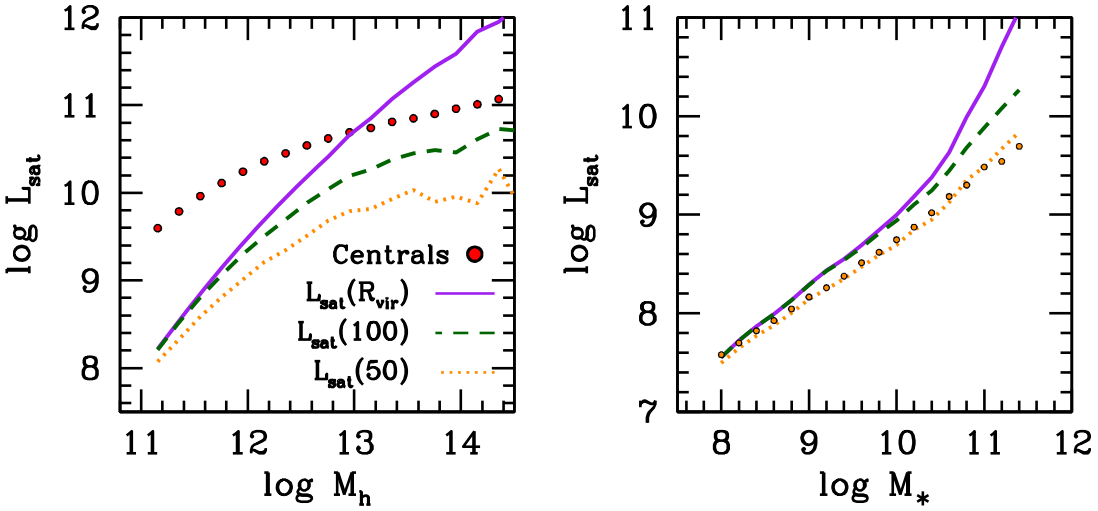

Figure 1. Left-hand panel: dependence of $L_{\rm sat}$ on $M_{\rm h}$ predicted by abundance matching applied to haloes in the C125 simulation. $L_{\rm sat}(R_{\rm vir})$ means all satellites within the virial radius, while 100 and 50 refer to $L_{\rm sat}$ within projected apertures, in h^{-1} kpc. For comparison, the luminosity of central galaxies is shown with the filled circles. At $M_{\rm h} \lesssim 10^{12}$, the luminosity of the central galaxy dominates the total luminosity within the halo. This is why using $L_{\rm sat}$ is a far more sensitive diagnostic of the dark halo than using $L_{\rm sat} + L_{\rm cen}$. Right-hand panel: same as left-hand panel, but now binning the object by M_* rather than $M_{\rm h}$. The filled circles show the results of a model in which the scatter of M_* at fixed $M_{\rm h}$ is perfectly correlated with $z_{1/2}$, the formation time of the halo. Although $L_{\rm sat}$ correlates with $z_{1/2}$, this does not impact results when binning in M_*.

higher density regions. Here were parametrize halo formation time as the redshift at which the halo reaches half its present-day mass, $z_{1/2}$.

Another halo property that correlates with formation history is concentration, c – early-forming haloes have higher concentrations. For aperture measurements of $L_{\rm sat}$ that are significantly smaller than $R_{\rm vir}$, we must determine the impact c has on estimates of $L_{\rm sat}$ at fixed halo mass.

Fig. 2 shows the dependence of $L_{\rm sat}$ on $z_{1/2}$ for three different bins of $M_{\rm h}$. The distribution of $z_{1/2}$ is shown with the histogram at the bottom of each panel. The filled circles show $L_{\rm sat}(R_{\rm vir})$, while the solid curve shows $L_{\rm sat}(50)$. Both show a significant dependence of $L_{\rm sat}$ on $z_{1/2}$, with younger haloes containing more satellites. Concentration is highly correlated with $z_{1/2}$, as shown by the blue squares in each panel. For these results, the y-axis is now $\log c/\langle c\rangle$. Younger haloes have lower concentrations, while older haloes have higher concentrations, in agreement with previous results (e.g. Wechsler et al. 2002; Giocoli, Tormen & Sheth 2012). The similarity between the circles and curves shows that concentration does not alter the relationship between $z_{1/2}$ and $L_{\rm sat}$, even when only measuring $L_{\rm sat}$ in the inner parts of the halo.

Haloes have myriad properties, most of which are correlated to some degree. In contrast to our direct approach in Fig. 2, machine learning has been used to investigate the full multidimensional dependence of halo and group properties on clustering (Han et al. 2019; Man et al. 2019). Incorporating $L_{\rm sat}$ into this approach may further shed light on how to connect observables to halo properties.

3.4 Dependence of $L_{\rm sat}$ on large-scale environment

The results of Fig. 2 indicate that $L_{\rm sat}$ will correlate with large-scale density at fixed $M_{\rm h}$ and M_*. This is shown explicitly in Fig. 3, in which we plot the clustering amplitude of galaxies relative to the clustering of dark matter on large scales. Different curves show the top and bottom quartiles of $L_{\rm sat}$ at fixed M_*. The lowest quartile in $L_{\rm sat}$ has significantly higher clustering than the highest quartile. However, the difference between the bias values is not as large as it would be if the galaxies were partitioned by $z_{1/2}$ directly. This is due to the fact that a large fraction of haloes, especially at low M_*, have no satellites brighter than $M_r - 5\log h = -14$. Thus, the trend of clustering with $z_{1/2}$ is lost for the fainter part of the $L_{\rm sat}$ distribution. For comparison, we show the prediction for the bias of the faintest quartile assuming a lower magnitude limit of -10 rather than -14. There are still many empty haloes, but the difference in clustering amplitudes is larger by nearly a factor of 2.

3.5 Assembly bias and $L_{\rm sat}$ scaling with M_*

Regardless of the degree to which $L_{\rm sat}$ correlates with $z_{1/2}$, if halo formation history does not correlate with galaxy properties, then it will not bias any results obtained through the $L_{\rm sat}$ method – $L_{\rm sat}$ would still be a true proxy for $M_{\rm h}$ if primary galaxies are stacked according to their observable properties. In detail this is likely to be not exactly true – halo abundance matching models based on $V_{\rm peak}$ or other formation-dependent halo quantities, rather than halo mass, are a better match[3] to observed galaxy clustering (Reddick et al. 2013; Lehmann et al. 2017; Zentner et al. 2019), and $V_{\rm peak}$ correlates weakly with large-scale environment at fixed halo mass, amplifying halo clustering by a few per cent (Walsh & Tinker 2019).

The right-hand panel of Fig. 1 considers the impact that galaxy assembly bias may have on how $L_{\rm sat}$ scales with M_*. If, for example, M_* correlates with $z_{1/2}$ at fixed $M_{\rm h}$, then what is the impact on the $L_{\rm sat}$–M_* correlation? To test this idea, we use the conditional abundance matching framework of Hearin & Watson (2013). After determining the mean M_*–$M_{\rm h}$ relation through our parametrized SHMR, but before adding scatter to the central galaxy M_*, we bin all host haloes in narrow bins of $M_{\rm h}$. In each bin, haloes are rank ordered by $z_{1/2}$

[3]We note that these conclusions are based on the standard χ^2 statistic to discriminate between models. Other studies (Sinha et al. 2018; Vakili & Hahn 2019) have found that extra parameters above halo mass are not statistically preferred when using more sophisticated statistical analyses. Thus, the observational situation is not fully settled.

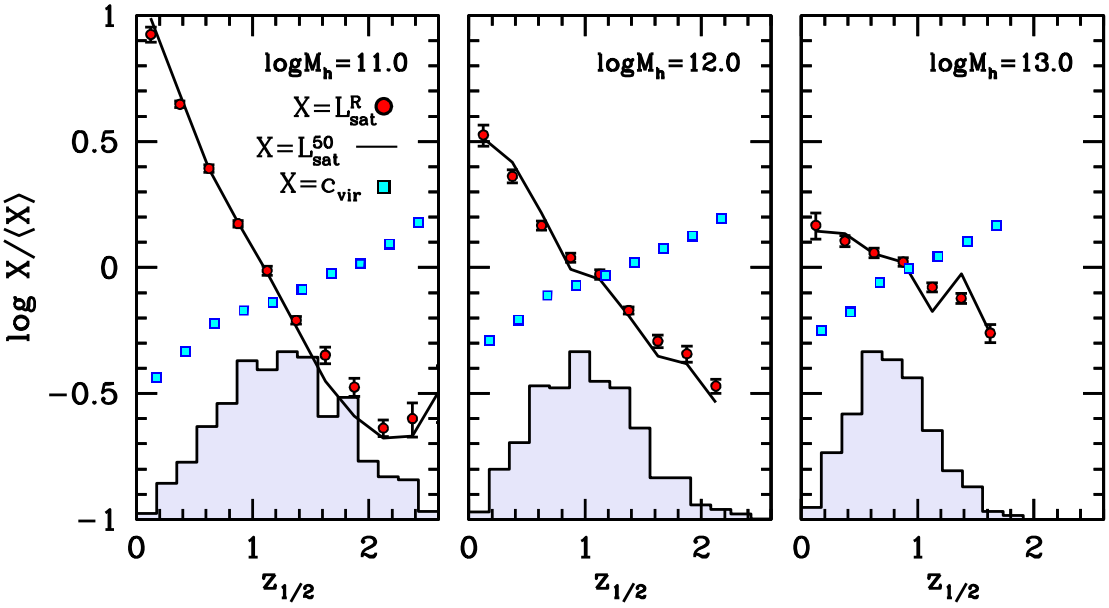

Figure 2. Dependence of halo properties on formation time, for different halo masses. In each panel, red circles show the correlation between $L_{\rm sat}(R_{\rm vir})$ and $z_{1/2}$ for haloes of different masses. The distribution of $z_{1/2}$ is shown with the histogram at the bottom of each panel. Blue squares show the correlation between c and $z_{1/2}$. The solid curves show $L_{\rm sat}(50)$, which is nearly the same as the results for $R_{\rm vir}$. This indicates that concentration c does not alter the relationship between $z_{1/2}$ and $L_{\rm sat}$, regardless of the aperture used. All results are based on the C125 simulation.

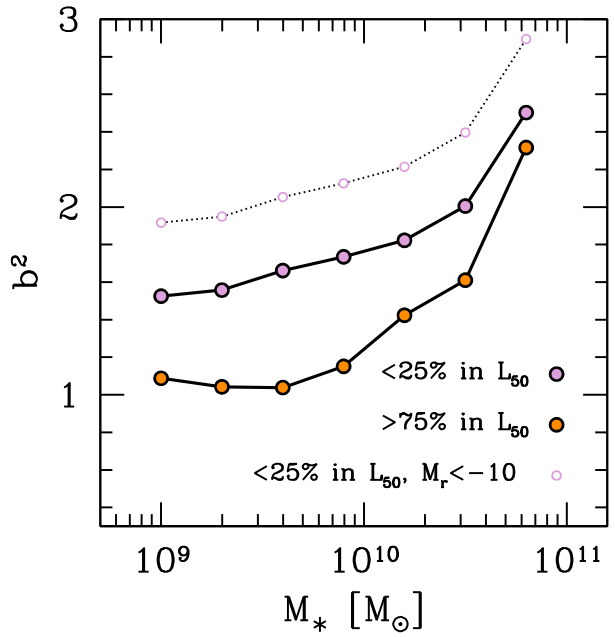

Figure 3. Clustering amplitude, expressed as b^2, as a function of M_*. Results here are based on the C250 simulation, combined with abundance matching models. The 'bias factor' b is defined as the ratio of the correlation function of haloes to that of matter $(\xi_{\rm h}/\xi_{\rm m})^{1/2}$ at $r > 10\,h^{-1}$ Mpc. Different coloured circles indicate different quartiles in $L_{\rm sat}$ at fixed M_*. Central galaxies in haloes with lower amounts of satellite light are more strongly clustered because these haloes have early formation times. The difference between the clustering amplitude of the upper and lower quartiles is attenuated by the fact that many of the haloes have no satellites about our fiducial magnitude threshold. Extending that threshold down to $M_r - 5\log h = -10$ from -14 further separates the quartiles because a smaller fraction is empty.

and matched to the *residual* with respect to the mean stellar mass, ΔM_*. The scatter about the mean is still a lognormal, but now M_* is correlated with $z_{1/2}$ at fixed $M_{\rm h}$. This yields a galaxy–halo connection very similar to that seen in both hydrodynamic models (Matthee et al. 2017) and semi-analytic models (Tojeiro et al. 2017), in which older haloes have more massive galaxies at fixed M_*.

To estimate the maximal possible effect, we assume a 1:1 correlation between $z_{1/2}$ and ΔM_* with no scatter. The results of this model for $L^{50}_{\rm sat}$ are shown in the right-hand panel of Fig. 1 with the orange circles. Even with the maximal assembly bias model, there is a negligible impact on $L_{\rm sat}$ scaling. This result would be quite different if we were binning on a secondary galaxy property at fixed M_* – properties such as effective radius or galaxy velocity dispersion – and these properties correlated with $z_{1/2}$. We will explore this possibility in a future work.

3.6 Distinguishing $M_{\rm h}$ from $z_{1/2}$

Fig. 4 shows an example of how to distinguish between the impact of $M_{\rm h}$ and $z_{1/2}$ on $L_{\rm sat}$. In the left-hand panel, two models are shown, constructed on the C250 simulation, in which $L_{\rm sat}$ correlates with a hypothetical secondary galaxy, X, at fixed M_* for central galaxies. The galaxy stellar mass is $\sim 10^{10}$ M$_\odot$, so the halo mass is $\sim 10^{11.8}$ M$_\odot$. For one model, X is correlated with $M_{\rm h}$ at fixed M_*, thus $L_{\rm sat}$ increases with X, even though stellar mass is held constant. For the second model, X anticorrelates with $z_{1/2}$ at fixed M_*. Thus, higher values of property X correlate with younger haloes, yielding a correlation between X and $L_{\rm sat}$ that is consistent between the two models.

In the right-hand panel of Fig. 4, we show the predictions these two models make for correlations between X and large-scale environment, $\rho/\bar{\rho}$. Here, ρ is measured in the same way as it would in the redshift survey data – measuring the density in a galaxy density field, in redshift space, in 10 Mpc top-hat spheres. This is consistent with the approach to measuring large-sale environment taken in Tinker et al. (2017, 2018b). The model in which X anticorrelates with $z_{1/2}$ is

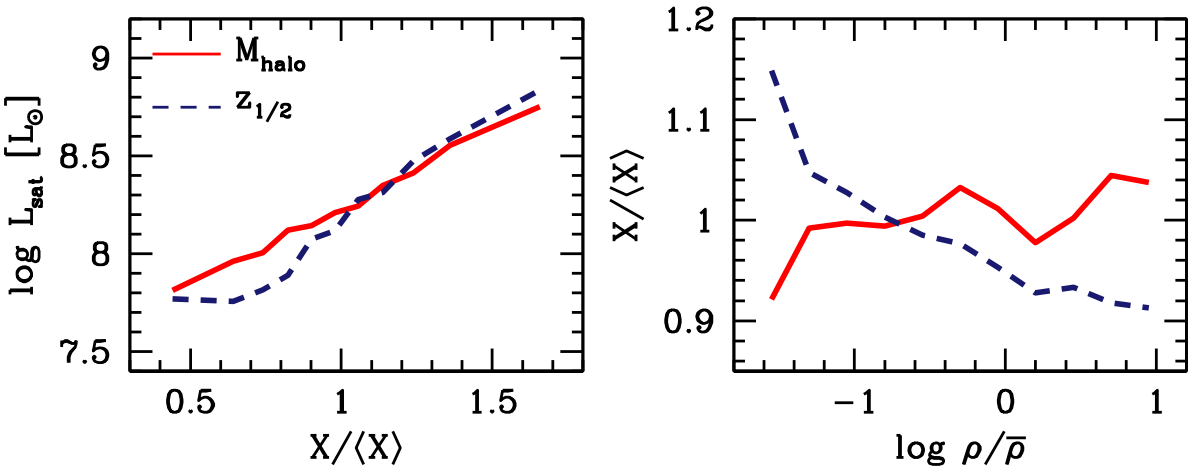

Figure 4. Left-hand panel: two theoretical models in which a hypothetical galaxy secondary property, X, correlates with the host halo at fixed M_*. The solid red curve is a model in which X correlates with $M_{\rm h}$, yielding a correlation between $L_{\rm sat}$ and X at fixed stellar mass. The dashed blue line is a model in which X anticorrelates with $z_{1/2}$ at fixed M_*, yielding a similar correlation. These two models can be distinguished in the right-hand panel, which shows the correlation between X and large-scale density ρ. The $z_{1/2}$ model shows a correlation between X and ρ, due to the fact that early-forming haloes reside in low-density regions. In the $M_{\rm h}$ model, the parameter X shows no correlation with density. All results are from the C250 simulation.

an example of galaxy assembly bias. There is a clear anticorrelation between X and ρ – younger haloes reside in underdensities, thus galaxies with higher values of X also are found in underdensities. The model in which X correlates with $M_{\rm h}$ shows no correlation with ρ – at the halo masses probed, there is little correlation between $M_{\rm h}$ and ρ. At higher values of M_*, the host halo masses will eventually get large enough such that a trend of X with ρ will be produced. But this trend will have opposite sign to the $z_{1/2}$ model – a positive correlation between X and $M_{\rm h}$ yields a positive correlation between X and ρ at $M_{\rm h}$ significantly larger than 10^{12} M$_\odot$.

Although the amplitude of the trend produced by the $z_{1/2}$ model is only ∼20 per cent, this amplitude is detectable in current SDSS data. Trends of this amplitude have been detected in star formation rate and Sersic index, while correlations with ρ can be ruled out in other parameters, such as galaxy colour or quenched fraction (Tinker et al. 2017, 2018b).

4 TESTING THE SELECTION OF CENTRAL GALAXIES

A critical aspect of this approach is the ability to select a pure sample of central galaxies. In this paper, we test two similar but complementary approaches to identify central galaxies within a spectroscopic galaxy sample: a galaxy group finder and a 'central galaxy finder,' as specified in Section 2. Although the halo masses estimated by the group finder can be suspect, especially when breaking the group sample into red and blue central galaxies, one of the group finder's strengths is identifying which galaxies are central and which are satellites. The code yields a quantity labelled $P_{\rm sat}$, which roughly corresponds to the probability of a given galaxy being a satellite in a larger halo. To characterize an entire galaxy population, splitting at $P_{\rm sat} = 0.5$ classifies all galaxies as one or the other. This breakpoint, however, does introduce some impurities into the sample of central galaxies of roughly 10–15 per cent (Tinker et al. 2011). Defining a sample of 'pure' central galaxies, with $P_{\rm sat} < 0.1$, removes the majority of impurities with only a small decrease in the completeness of the sample. Defined in this way, pure centrals are not biased in terms of the distribution of environments in which central galaxies are found (Tinker et al. 2017).

The second method, the central galaxy finder, is more flexible but less robust. We describe the method in Appendix A. Briefly, the method is quite similar to the group finder, but makes no attempt to find true groups. The algorithm first assigns halo masses to galaxy masses using *inverse* abundance matching – in contrast to using abundance matching to assign M_* to haloes in simulations. We use the tabulated SHMRs of Behroozi, Wechsler & Conroy (2013d) rather than using actual galaxy counts. Thus, the method does not require that a galaxy sample be volume limited, or contain a large number of galaxies. To classify whether a galaxy is a central, all that is required is to know the masses of its nearest neighbours. Using the properties of haloes – their velocity dispersions and density profiles – the code determines the probability that a given galaxy is a central galaxy in its halo. We refer to this quantity as $P_{\rm cen}$.

4.1 Bias on the $L_{\rm sat}$–M_* relation

Fig. 5 shows the impact of impurities in the central galaxy samples has on $L_{\rm sat}^{50}$. To perform this test, we created mock galaxy distributions using the C250 box that match the PCA stellar mass function. Each galaxy in the mock is assigned both an $L_{\rm sat}^{50}$ value and total luminosity of interloper, or 'background' galaxies we term $L_{\rm BG}$. Thus the total light within the aperture is $L_{\rm tot} = L_{\rm sat}^{50} + L_{\rm BG}$. These are chosen to match the distributions of $L_{\rm tot}$ and $L_{\rm BG}$ seen in the data, which we describe in Section 5. Results in Fig. 5 represent the mean of 100 mock realizations. In each realization, the quantity that varies is the assignment of $L_{\rm sat}^{50}$, which is drawn randomly from the distribution of $L_{\rm tot}$ and $L_{\rm BG}$ in the SDSS data.

Central and satellite galaxies follow different distributions of $L_{\rm sat}^{50}$. The left-hand panel of Fig. 5 shows the input values of $L_{\rm sat}^{50}$ as a function of M_* of the mock. At a given M_*, the $L_{\rm sat}$ values for satellite galaxies are roughly an order of magnitude higher than for centrals. Thus, impurities in the central sample can bias the measurements if the purity goes significantly below unity.

To construct a mock galaxy survey suitable for testing the group finder, we first convert the Cartesian positions and velocities of the cubic mock from C250 to RA, Dec., and redshift assuming one corner of the box as the observer. The resulting mock is an octant of the full sky, volume limited down to $M_* = 10^9$ M$_\odot$, with a maximum

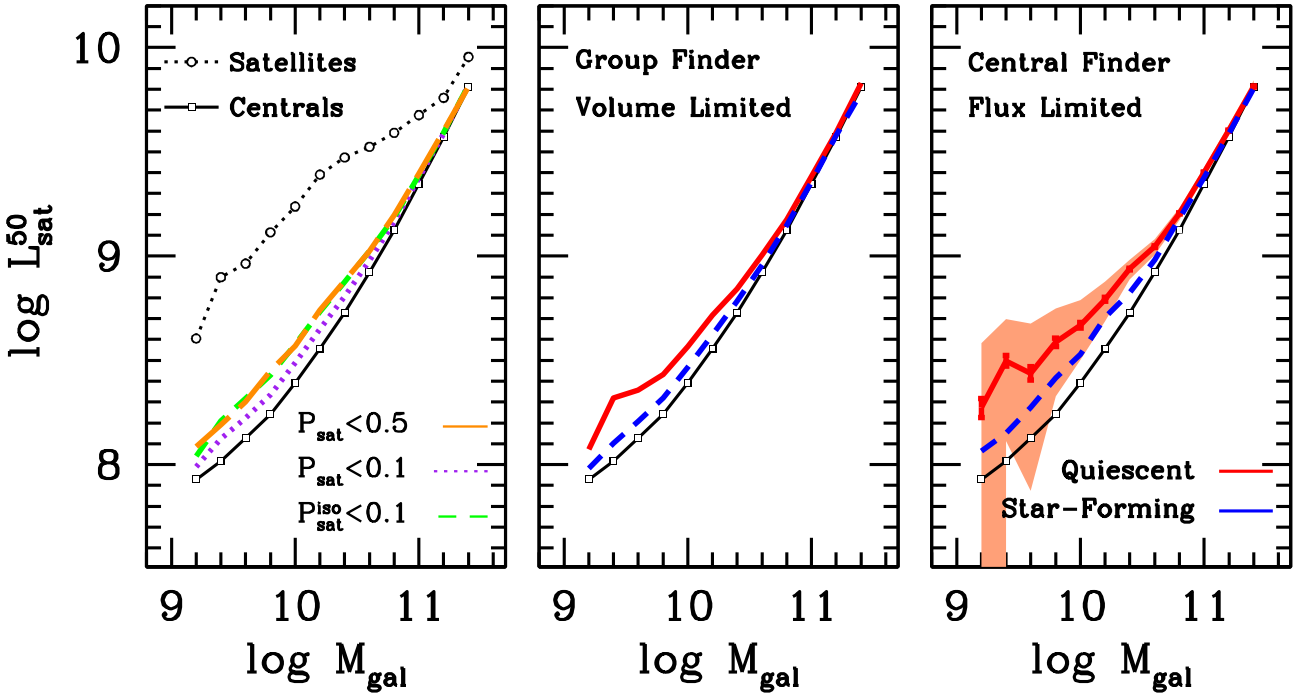

Figure 5. Impact of the selection of central galaxies of the values of $L_{\rm sat}^{50}$. The open squares connected by solid curve show the input value of $L_{\rm sat}^{50}$ around central galaxies. The open circles connected by the dashed black curve show the values of $L_{\rm sat}^{50}$ if satellite galaxies are taken to be the primary galaxies. Satellites reside in higher mass haloes, thus their $L_{\rm sat}^{50}$ values are larger than those for central galaxies of the same M_*. Left-hand panel: $L_{\rm sat}^{50}$ values for the full sample of central galaxies constructed from mock galaxy distributions. The long-dashed and dotted curves show the results from applying the galaxy group finder to a volume-limited mock sample of galaxies. The long-dash line uses the full sample of central galaxies ($P_{\rm sat} < 0.5$), while the dotted line shows results for the pure sample of central galaxies ($P_{\rm sat} < 0.1$). The short-dashed line shows the results of applying the central galaxy finder to a flux-limited mock galaxy catalogue, using $P_{\rm cen} > 0.9$ to define the sample. Middle panel: $L_{\rm sat}^{50}$ values obtained from applying the group finder to a volume-limited mock galaxy sample, now broken into central galaxies that are star forming and quiescent. See text for details. The central galaxy sample is defined as $P_{\rm sat} < 0.1$. The input model assumes that all galaxies at fixed M_* live in the same mass haloes, so the any difference in the measured $L_{\rm sat}^{50}$ values represents bias incurred by the group finding algorithm. Right-hand panel: same as the middle panel, only now the mock galaxy sample is flux limited, matching the $n(z)$ of the SDSS data, and central galaxies are identified using the central galaxy finder with $P_{\rm cen} > 0.9$. The errors on the results for quiescent galaxies represent the error in the mean of 100 realizations of the mock. The shaded area represents the dispersion.

redshift of $z = 0.08$. The group finder is then applied to the mock. We use the group finder results to make two central galaxy samples: one with $P_{\rm sat} < 0.5$ and a pure sample with $P_{\rm sat} < 0.1$. The results are shown with the long-dash and dotted lines in the left-hand panel. Above $M_* = 10^{10.5}$ M$_\odot$, the bias induced by the impurities in the sample is negligible. At lower masses, the $P_{\rm sat} < 0.5$ sample yields a bias of $\sim$0.15 dex. The pure sample of $P_{\rm sat} < 0.1$ yields a bias roughly half as large, at $\sim$0.08 dex.

To apply the central galaxy finder to the mock, we modify the procedure to construct a flux-limited sample, rather than a volume-limited sample. First, we repeat the C250 cubic mock eight times to extend the maximum redshift limit to $z = 0.16$. The mock is divided into narrow bins of $\log M_*$, which are subsampled match the observed redshift distributions $N(z|M_*)$ in the full flux-limited SDSS catalogue, in units of [dz^{-1} deg^{-2}]. We do this process separately for star-forming and quiescent galaxies, because quiescent galaxies are fainter at fixed M_*, thus they have distinct redshift distributions.

We apply the central galaxy finder to the flux-limited mock catalogue, using $P_{\rm cen} > 0.9$ to define the sample of central galaxies. The $L_{\rm sat}^{50}$ results, shown in the left-hand panel of Fig. 5 as well, are nearly indistinguishable from the full $P_{\rm sat} < 0.5$ sample from the group finder. Thus, the central galaxy sample from the flux-limited catalogue has slightly more bias in it than the pure sample from the volume-limited catalogue, but the flexibility increased statistics afforded through the flux-limited catalogue make it possible to perform fine-grained, multidimensional binning on $L_{\rm sat}$ results.

4.2 Bias on $L_{\rm sat}$ in subpopulations

The true concern of impurities in the central sample is not in biasing the results of $L_{\rm sat}^{50}$ for the full sample, but rather creating a differential bias for subclasses of galaxies that have a higher fraction of satellite galaxies in them. The prime example of such a sample is quiescent galaxies – quiescent galaxies have a significantly higher satellite fraction than star-forming galaxies at all stellar masses (e.g. Weinmann et al. 2006; Wetzel, Tinker & Conroy 2012; Tinker et al. 2013). In the mock, we constrain the quenched fraction to match that seen in the group finder as a function of stellar mass. We do this separately for true centrals and true satellites in the mock, as they have different quenched fractions. We assume that star-forming and quiescent central galaxies of the same M_* live in haloes of the same $M_{\rm h}$. This choice is driven less by results in the literature (see the wide disparity of results in Wechsler & Tinker 2018 described in Section 1) than it is to make it straightforward to identify any biases imparted by impurities in the central sample: the input model yields $L_{\rm sat}^{50}$–M_* correlations that are identical for star-forming and quiescent galaxies. Thus any differences in the final results are a consequence of impurities in the central galaxy sample.

The middle panel of Fig. 5 shows the results of applying the group finder to the volume-limited mock, and then dividing up the sample by star-forming and quiescent galaxies. Results here are for $P_{\rm sat} < 0.1$. The results for star-forming galaxies are nearly identical to those for the full sample in the left-hand panel, but the quiescent galaxies have slightly higher $L^{50}_{\rm sat}$ values, roughly twice the bias (in dex).

The right-hand panel of Fig. 5 shows the results of applying the central galaxy finder to the flux-limited mock catalogue. The amplitudes of the biases are larger than that seen in the results from the group catalogue, but the results are consistent, with the bias in $\log L^{50}_{\rm sat}$ for quiescent galaxies being twice as large as that seen in the star-forming central galaxies.

5 APPLICATION TO DATA

In this section, we will discuss how we ameliorate observational systematics that may influence the measured values of $L^{R}_{\rm sat}$ and $L^{50}_{\rm sat}$. We focus here on variable survey depth for $L^{R}_{\rm sat}$, unbiased estimation of the background galaxy counts for $L^{50}_{\rm sat}$, and eliminating redshift dependence of the $L^{50}_{\rm sat}$ measurements. The results presented in this section will show measurements conducted on actual Legacy Surveys data.

5.1 Measuring the conditional luminosity function within dark matter haloes

The conditional luminosity function (CLF) is defined as the luminosity function of satellite galaxies, conditioned on the mass of the host halo, $\Phi_{\rm sat}(M_r|M_{\rm h})$. The depth of the DLIS data varies across the sky, thus when measuring the CLFs, we track the limiting magnitude for each halo depending on its angular position and redshift. For each primary galaxy, we assume all photometric galaxies along the line of sight are at the redshift of the primary galaxy. For each halo, we define the limiting absolute magnitude $M_{r,\rm lim} = m_{r,\rm lim} + \mu$, where μ is the distance modulus at the redshift of that halo and $m_{r,\rm lim}$ is the limiting magnitude at that location in the survey. The luminosity function is defined as

$$\Phi_{\rm sat}(M_r|M_{\rm h})\,{\rm d}M_r = \frac{N_{\rm gal}(M_r)}{N_{\rm h}(M_r)}, \quad (1)$$

where $N_{\rm h}(M_r)$ is the number of haloes that have a $M_{r,\rm lim} > M_r$, and $N_{\rm gal}(M_r)$ is the number of galaxies above the background at magnitude M_r, expressed as

$$N_{\rm gal}(M_r) = N_{\rm tot}(M_r) - f_{\rm A} N_{\rm BG}(M_r), \quad (2)$$

where $N_{\rm tot}$ is the total number of galaxies at M_r within the aperture centred on the SDSS galaxy, and $N_{\rm BG}$ is the number of background galaxies at that magnitude. This quantity is measured from an annulus around the halo. This is the same approach as taken in Hansen et al. (2009) and Tal et al. (2012) for estimating the background around massive objects, such as clusters or luminous red galaxies. Our fiducial choice for the annuli radii are $R = [R_{\rm vir}, 3R_{\rm vir}]$, but in practice we find that the exact choice of annuli boundaries has negligible impact on the results. The factor $f_{\rm A}$ is to account for any differences in the area used to estimate the two N values. For example, if we use annuli with $R = [R_{\rm vir}, 3R_{\rm vir}]$, $f_{\rm A}$ for $L^{R}_{\rm sat}$ is $1/(9-1) = 1/8$.

5.2 CLF recovery in simulations

Fig. 6 shows the results of applying our method of measuring the CLF on simulations. The mock galaxy catalogues we use are the Buzzard mocks (DeRose et al. 2019). These simulations were utilized by the Dark Energy Survey (DES) to test the pipeline and quantify systematic errors in the cosmological analysis (MacCrann et al. 2018). The Buzzard mock galaxy distributions have large volume, subtending roughly 10 000 deg^2 when projected on to the sky. The mocks also incorporate galaxies faint enough to match the flux limit in DES imaging, which is significantly deeper than DLIS data. The mocks are tuned to match the evolution of the luminosity function from $z = 0$ to $z = 1$, galaxy clustering as a function of luminosity, and the observed colour–density relation at $z = 0$. In order to create such a large-volume mock, the dark matter simulation on which the mock is built only resolves haloes down to $M_{\rm h} \approx 10^{12.5}$ M$_\odot$, and the galaxies that would be contained in lower mass haloes are placed by sampling the dark matter density field in order match all of the observational statistics mentioned above.

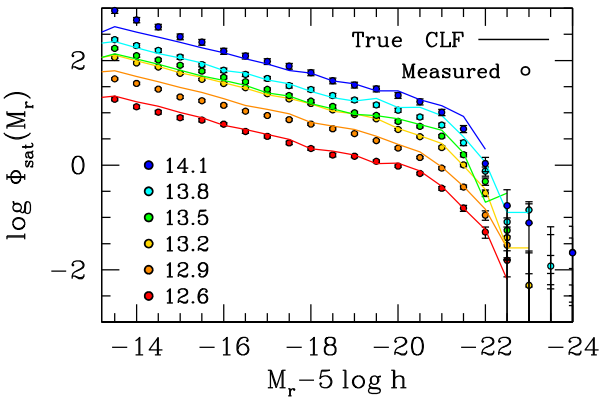

Figure 6. Recovery of CLFs in simulations. Here we present results using the Buzzard mock galaxy samples (DeRose et al. 2019). The Buzzard simulations are constructed to match observational statistics of an $r < 24$ photometric sample of galaxies. The method of measuring $\Phi_{\rm sat}$ uses annuli around each primary to estimate the background. We find minimal dependence of results on the exact choice of the annuli radii, with the exception of the highest mass bin, $M_{\rm h} > 10^{14}$ M$_\odot$.

To mimic our analysis of $L^{R}_{\rm sat}$ on DLIS data, we restrict the Buzzard mock to galaxies with $r < 24$, and construct a volume-limited sample of primary galaxies from all resolved haloes within $z = 0.1$. Here we assume perfect knowledge of the true sample of central galaxies, thus this test isolates the method of stacking haloes and background subtraction. As with the DLIS data, we measure the background using annuli around each primary object. Results are shown in Fig. 6 in bins of $\log M_{\rm h}$. The recovered CLFs are in excellent agreement with the mock inputs, including how the amplitude of $\Phi_{\rm sat}$ scales with $M_{\rm h}$ and the magnitude of the cut-off in each luminosity function. This test demonstrates that the interlopers present in both the annuli and the apertures around the primary galaxies cancel each other, allowing robust recovery of $\Phi_{\rm sat}$. To measure $L^{R}_{\rm sat}$, we integrate the measured values of $\Phi_{\rm sat}(M_r|M_{\rm h})$, weighted by the luminosity at each bin of M_r.

5.3 Measuring $L^{50}_{\rm sat}$

As discussed above, when measuring $L_{\rm sat}$ around galaxies – rather than haloes – we wish to impart no prior on $M_{\rm h}$ in the measurements. Thus we choose $L^{50}_{\rm sat}$ as our observational quantity around galaxies. To measure $L^{50}_{\rm sat}$ we first estimate the CLF individually for each galaxy, but restricting the galaxy counts to be within the 50 h^{-1} kpc aperture centred on the primary galaxy. Our method for measuring $L^{50}_{\rm sat}$ differs from $L^{R}_{\rm sat}$ in our approach to estimating $N_{\rm BG}(M_r)$ in equation (2), which we discuss in the proceeding section. After making the measurement of the 50 h^{-1}kpc CLF, we integrate $\Phi(M_r)$

from $M_r^{\rm lim}$:

$$L_{\rm sat} = \sum_{M_r^{\rm low}}^{M_r^{\rm hi}} 10^{-0.4(M_r - M_{r,\odot})} \Phi(M_r) \Delta M_r, \quad (3)$$

where $M_{r,\odot} = 4.65$ and we use the centre of the bin for the value of M_r. In equation (3), the limits of the summation are set to maximize the signal-to-noise ratio of the $L_{\rm sat}$ measurement. In practice, background galaxies at low redshift (but below the magnitude limit of SDSS spectroscopy) can cause large fluctuations in $L_{\rm sat}$ when converting to absolute magnitude. Thus we enforce a bright limit based on the stellar mass of the central galaxy, such that

$$M_r^{\rm hi} = -21 - 2\,(\log M_* - 10). \quad (4)$$

Essentially, equation (4) enforces a limit that satellite galaxies cannot be brighter than the central galaxy itself. For $M_r^{\rm low}$ we chose $M_r - 5\log h = -14$. At fainter magnitudes, we find that the luminosity function of all galaxies within the aperture centred on the central galaxy falls *below* the luminosity function of the background at fainter magnitudes (i.e. equation 2 would be negative). The presence of bright foreground galaxies – i.e. the SDSS centrals and any satellites within the halo – reduces the efficiency of finding faint extended objects in the DLIS imagery. We choose -14 as our faint limit because it yields a nearly volume-limited sample out to $z = 0.15$, which is the maximum redshift in our flux-limited sample.

5.4 Estimating the background for $L_{\rm sat}^{50}$

As mentioned above, we use a different approach to estimating the background surface density of galaxies when measuring $L_{\rm sat}^{50}$ than when measuring satellite luminosities within the entire halo. When using annuli at $\sim 3R_{\rm vir}$ to estimate the background, most of these annuli are impacted by the presence of interlopers – other SDSS galaxies – projected along the line of sight. Thus, the estimate of $L_{\rm BG}$ is biased due to the presence of haloes that contain an enhanced number of secondary galaxies within them.

When estimating $L_{\rm sat}^{R}$, using annuli does not actually bias the value of $L_{\rm sat}$ obtained – the apertures subtended by the halo $R_{\rm vir}$ suffers from the same contamination at the annuli themselves (as demonstrated in the comparison to the Buzzard mocks in Fig. 6). However, when estimating $L_{\rm sat}^{50}$, the smaller aperture yields a much smaller fraction of $L_{\rm tot}$ measurements that have interlopers. Although it is possible to develop an estimator to robustly account for the presence of interlopers in the background (i.e. Masjedi et al. 2006), these primary objects can be successfully removed from the sample of central galaxies without significantly impacting the statistics of the overall sample. Therefore, to measure an unbiased $L_{\rm sat}^{50}$ we require an estimate of $L_{\rm BG}$ that is also unaffected by interlopers.

Fig. 7 shows how interlopers impact the value of $L_{\rm BG}$. The y-axis shows the mean surface brightness of background galaxies in annuli around primary galaxies with $M_* = 10^{10.5}$ M$_\odot$. All these annuli contain interlopers within them, thus for each annulus we measure the total stellar mass of all interlopers within the annulus. The x-axis, labelled 'interloper limit,' refers to the fraction of annuli used in calculating the mean surface brightness of background galaxies. The annuli are rank ordered by the total stellar mass of the interlopers. An interloper limit of 0.6 means than the lower 60 per cent of annuli are used to calculate the mean background surface brightness. Thus, the background surface brightness decreases monotonically with decreasing interloper limit.

The red circle at zero on the x-axis is an estimate of the background from randomly chosen areas within the footprint that do not contain interlopers. Specifically, this value is the average of $\sim$30 000 apertures with angular radius 35.2 arcsec, which is 50 h^{-1} kpc at $z = 0.05$. The locations of these points were chosen from random locations within both the SDSS mask and the DLIS footprint, and then restricted to only include locations that are outside $R_{\rm vir}/2$ of the nearest SDSS galaxy. Extending this limit to only include locations that are fully outside any SDSS virial radius does not leave enough area for a robust estimation of the background. But, as can be seen from Fig. 7, using these random locations yields an estimate of the background that is a clear extrapolation of the annuli toward having no interlopers. In all measurements of $L_{\rm sat}^{50}$, we use the mean $\langle N_{\rm BG}(M_r)\rangle$ from these 30 000 apertures as our estimate of the background surface density of galaxies.

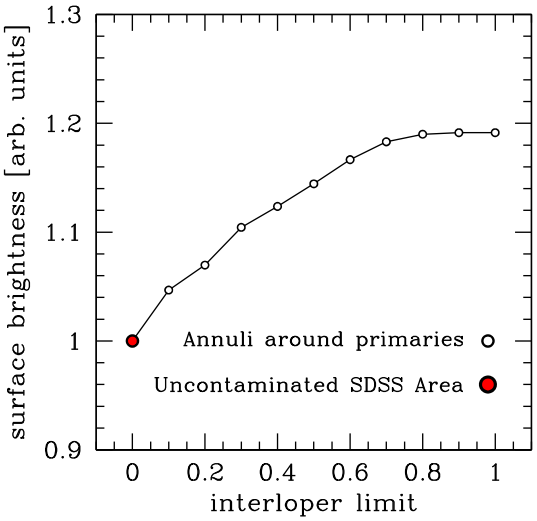

Figure 7. Open circles show the estimates of the surface brightness in background galaxies using annuli around SDSS primary galaxies. All annuli have some contamination by interlopers (other primary galaxies). The annuli are rank ordered by total stellar mass of SDSS galaxies within them, and thus 'interloper limit' indicates what fraction of annuli are used to calculate the background. The filled red circles at $x = 0$ is an estimate of the background from random locations that have no interlopers.

5.5 Redshift effects on $L_{\rm sat}^{50}$

We do not expect the galaxy–halo connection to change significantly over the redshift baseline of the SDSS. Thus, a good diagnostic for our method of estimating $L_{\rm sat}$ is to demonstrate that the quantity is independent of z. Fig. 8 shows results with DLIS data to demonstrate this fact.

Fig. 8 shows $L_{\rm sat}^{50}$ as a function of z for three different stellar mass bins. In each panel, we show three different methods of calculating $L_{\rm sat}^{50}$. Red points show measurements using annuli to estimate the background, and all primary galaxies are included regardless of whether there are interlopers within the 50 h^{-1} kpc aperture. Green points show the results for all primary galaxies when we use the uncontaminated random points for the estimate of $L_{\rm BG}$. Blue points – our fiducial method – show the results using the random points for $L_{\rm BG}$, but now removing primary objects with interlopers within the aperture. The shaded histogram at the bottom of each panel shows the redshift distribution for each sample of galaxies.

In each panel, the red points (annuli background) show significant redshift trend of $L_{\rm sat}^{50}$. For the lowest M_* bin, the values of $L_{\rm sat}^{50}$ become negative at the peak of the redshift distribution. Switching from annuli backgrounds to the random background avoids this problem,

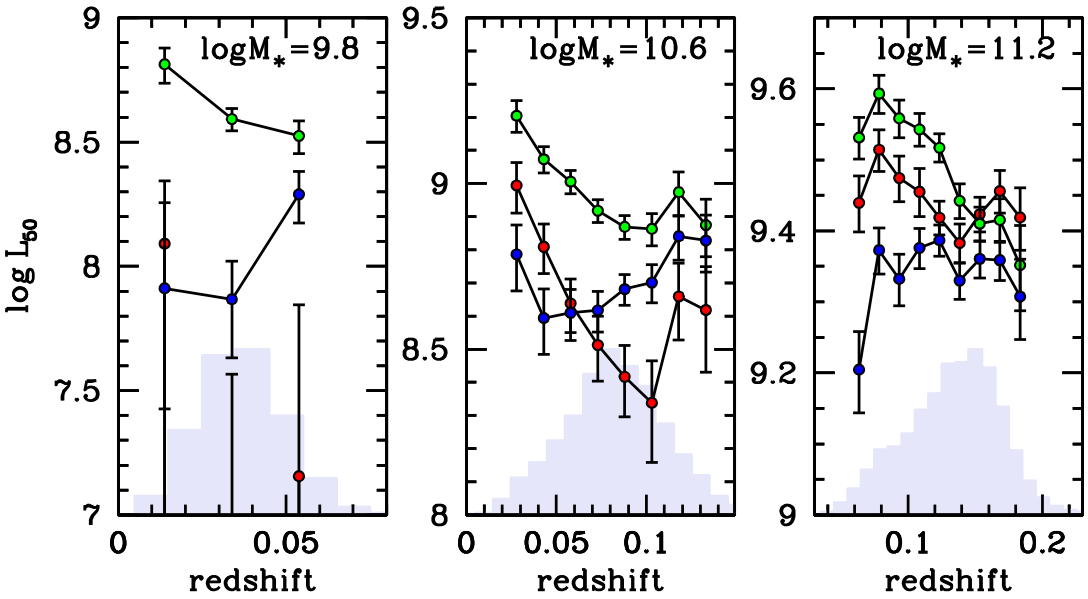

Figure 8. Results for SDSS primary galaxies: $L_{\rm sat}^{50}$ as a function of z for three different methods of measuring $L_{\rm sat}$. In each panel, the blue points indicate $L_{\rm sat}^{50}$ after removing primary galaxies that have interlopers inside the 50 h^{-1} kpc aperture, and the background luminosity of galaxies is taken from random points within the SDSS footprint (see text for details). This method produces $L_{\rm sat}^{50}$ results that are roughly constant with redshift, and thus we use it as our fiducial method. Red points indicate $L_{\rm sat}^{50}$ when using annuli around the primaries to estimate the background, and including all primaries that are within the DLIS footprint. This method yields a strong redshift trend, and even produces some negative values of $L_{\rm sat}^{50}$. Green points represent $L_{\rm sat}^{50}$ when using random points to estimate the background, but still include all primary galaxies. This method also produces a strong redshift trend. Results are shown for three different bins in M_*. Below $M_* = 10^{9.8}$ M$_\odot$, the data do not span a wide enough range in redshift to quantify and trend in z. The shaded histogram at the bottom of each panel shows the redshift distribution for galaxies in the M_* bin.

since now the estimate of $L_{\rm BG}$ is ∼20 per cent lower, but there is still a significant redshift trend with $L_{\rm sat}^{50}$. The blue points, where primary objects with interlopers are removed, yield $L_{\rm sat}^{50}$ values that are roughly independent of redshift. Errors on each data point are from bootstrap resampling on the sample of primary galaxies in each bin in redshift.

6 RESULTS FOR SDSS CENTRAL GALAXIES

In this section, we present our first results of measuring $L_{\rm sat}$ around SDSS galaxies. We focus on how our measurements of $L_{\rm sat}$ scale with $M_{\rm h}$ and M_*, and how they compare to our abundance matching predictions. The values of $M_{\rm h}$ are taken from the group catalogue, and thus should be taken as estimates only. Analysis of the $L_{\rm sat}$ results for subpopulations of galaxies, secondary galaxy properties, and correlations with environment will be presented in further work.

6.1 Projected radial profiles

Fig. 9 shows the surface number density of photometric galaxies, centred on primary identified through the galaxy group finder. In the left-hand panel, each colour indicates a different bin in $M_{\rm h}$. As halo mass increases, the amplitude of $\Sigma_{\rm gal}$ also increases, but as $R_{\rm p}$ decreases, each projected density profile hits a maximum value and then turns over at smaller scales. This is due to the presence of a bright central galaxy located at the centre of the halo. The location of this peak value increases with $M_{\rm h}$, as expected as the luminosity of the central galaxy increases as well. The thick blue curve in this panel shows the total projected number density of galaxies before background subtraction, for $M_{\rm h} = 10^{13.2}$ M$_\odot$. The value of the background can easily be seen, as $\Sigma_{\rm gal}$ reaches a horizontal asymptote well within the virial radius of the halo. However, this curve shows the same behaviour as the measurements for satellite galaxies, peaking at $R_{\rm p} \sim 15$ kpc h^{-1}. Inside this scale, the total $\Sigma_{\rm gal}$ falls below the asymptotic value of the background, demonstrating the presence of the central galaxy inhibits source detection at the centre of the halo.

The right-hand panel of Fig. 9 compares our measurements for $M_{\rm h} = 10^{11.1}$ M$_\odot$ haloes to several theoretical predictions. The solid red curve shows the abundance matching prediction from the C125 simulation. The thick dashed line shows a Navarro–Frenk–White (NFW) profile (Navarro, Frenk & White 1997) with concentration parameter of $c = 15$, which is consistent with the ΛCDM prediction for haloes of this mass (Macciò, Dutton & van den Bosch 2008). The blue dotted line is a modified NFW profile, with inner slope $\gamma = -2$ rather than -1, but with the same concentration parameter. The dot–dash line shows the Einasto profile fit to the distribution of subhaloes in the Aquarius project (Springel et al. 2008).

The abundance matching prediction is significantly shallower than the SDSS results or the NFW profiles, although it is in good agreement with the Aquarius results, which are significantly higher resolution. This difference is expected from numerical simulations of dark-matter-only substructure. Han et al. (2016) demonstrated that tidal stripping in collisionless simulations accounts for the differences between radial distributions of dark matter subhaloes and observed radial distributions of galaxies, which tend to be much steeper. There is additional observational evidence consistent with this claim as well: when constructing models of faint Milky Way satellite from collisionless N-body simulations, Nadler et al. (2019) note that 'orphans' – evolving satellite galaxies analytically after their host subhaloes are too disrupted to track within the simulation – are required to match the total observed number of satellites. Previous comparisons between the small-scale clustering of subhaloes and measured clustering of galaxies have shown good

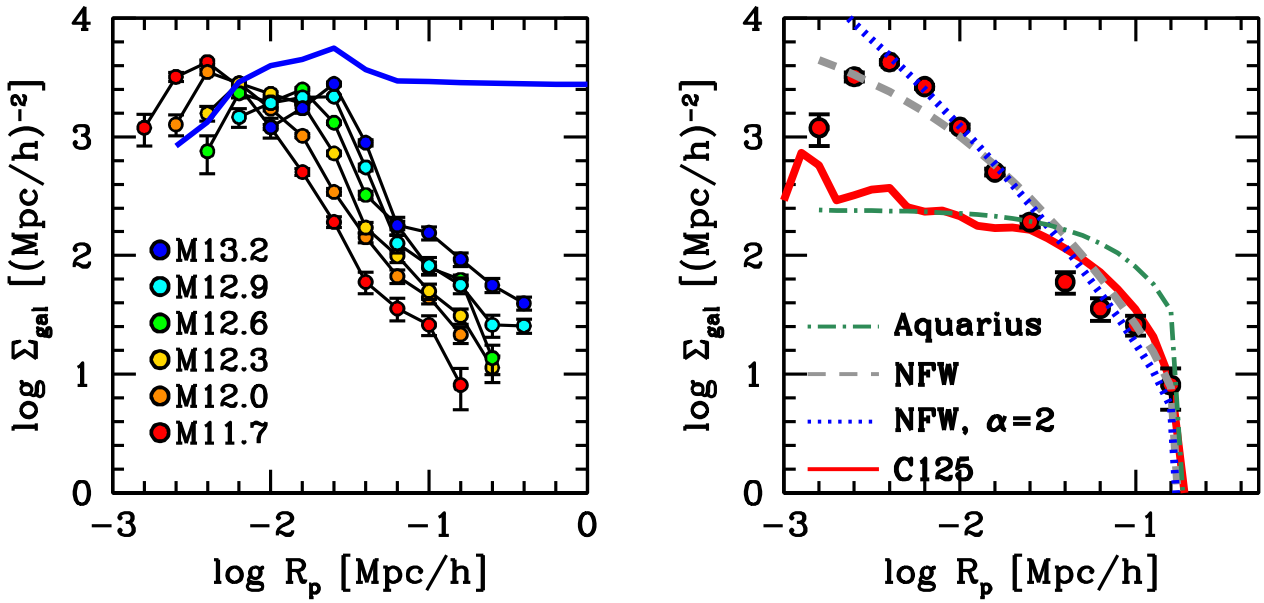

Figure 9. Left-hand panel: the excess surface density of galaxies above random, $\Sigma_{\rm gal}$, as a function of projected comoving separation from the group centre. To calculate distances, all galaxies are assumed to be at the redshift of the group. The key shows the value of $\log M_{\rm h}$. The solid blue curve shows the total projected number of densities of galaxies before background subtraction. All data are for galaxies brighter than $M_r - 5\log h = -14$. The turnover of all these data at small scales is due to the presence of a bright central galaxy in the group. Right-hand panel: $\Sigma_{\rm gal}$ for the $M_{\rm h} = 10^{11.7}$ bin compared to different theoretical models. The solid red curve shows the abundance matching result from the C125 simulation, which is consistent with the higher resolution Aquarius simulations. The dash and dotted curves show Navarro–Frenk–White (NFW) fits with concentration parameter $c = 15$, but with inner slopes of $\gamma = -1$ (standard NFW) and $\gamma = -2$, which is supported by small-scale clustering measurements of Watson et al. (2012).

agreement (e.g. Conroy, Wechsler & Kravtsov 2006; Reddick et al. 2013; Lehmann et al. 2017), but these comparisons are restricted to scales larger than those of interest here ($R_{\rm p} \gtrsim 0.2\,h^{-1}$ Mpc) and for brighter galaxy samples. Clustering measurements at these scales are largely insensitive to the values of halo concentration parameters used (Tinker et al. 2012). Watson et al. (2012) use cross-correlation techniques to measure galaxy clustering down to $R_{\rm p} \sim 0.01\ h^{-1}$ Mpc, finding evidence for steeper inner density profiles, consistent with the $\gamma = -2$ shown in Fig. 9, and concentration parameters roughly consistent with that predicted by collisionless N-body simulations.

Either NFW profile is a reasonable description of the data, relative to the subhalo results, but all $\Sigma_{\rm gal}$ measurements have a 'kink' at $R_{\rm p} \sim R_{\rm vir}/5$. This feature can be seen in the results for the lowest halo mass bin in the right-hand panel of Fig. 9, at $\log R_{\rm p} \approx -1.4$, where the SDSS measurements fall below the predictions for C125. This could be an artefact of the small amount of satellite galaxies leaking into the sample of central galaxies, or an observational systematic related to background subtraction. Further study is required for a sufficient explanation of this feature.

6.2 Conditional luminosity functions

Fig. 10 presents our measurements of the CLFs within the first and third volume-limited samples listed in Table 1. To avoid crowding, we split the CLFs into two panels. In each panel, the connected circles show the results from our method of stacking DLIS data and subtracting off the background contribution estimated in annuli around each halo. Errors are estimated by bootstrap resampling on the sample of central galaxies. The thick coloured curves show results from DLIS galaxies that are brighter than the SDSS flux limit, and thus have spectroscopic redshifts. For each DLIS galaxy with a redshift, we calculate the probability that is a satellite within a group in our SDSS group catalogues using the same procedure that the group catalogues were constructed in the first place (see appendix A in Tinker et al. 2011). These two independent methods of estimating $\Phi_{\rm sat}$ show excellent agreement in the regions they overlap. The spectroscopic sample is more efficient at removing extremely bright galaxies that are spuriously counted as satellites, but the overall amplitudes and Schechter function breakpoints agree between the two measurements.

Fig. 11 compares our CLF measurements to those of Lan et al. (2016). Their method of identifying central galaxies is similar with ours, as our group finder is based on the same Yang et al. (2005) algorithm. The differences in the analyses are in the central sample construction – we use volume-limited samples, while they use the flux-limited samples – the imaging data used to quantify the satellite galaxies, and a small difference in the cosmology adopted for the halo mass function and distance–redshift relation. None the less, the consistency between the two analyses is encouraging.

Fig. 12 compares the CLF measurements to abundance matching predictions. Here, the data points are a composite of the imaging results and the spectroscopic results. Overall, the agreement is reasonable. The overall scaling of $\Phi_{\rm sat}$ with halo mass is in good agreement (we will show integrated values of $L_{\rm sat}$ in the next section), and the abundance matching predictions match the observed CLFs at the high-luminosity end. The primary disagreement between the N-body prediction and the measurements is in the power-law slope at faint luminosities, which is shallower in the data than in the abundance matching model. The reason for this discrepancy is not immediately clear. It could that our extrapolation of our particular implementation of the abundance matching model to such low luminosities is no longer valid. We will address these questions in more detail in a subsequent paper.

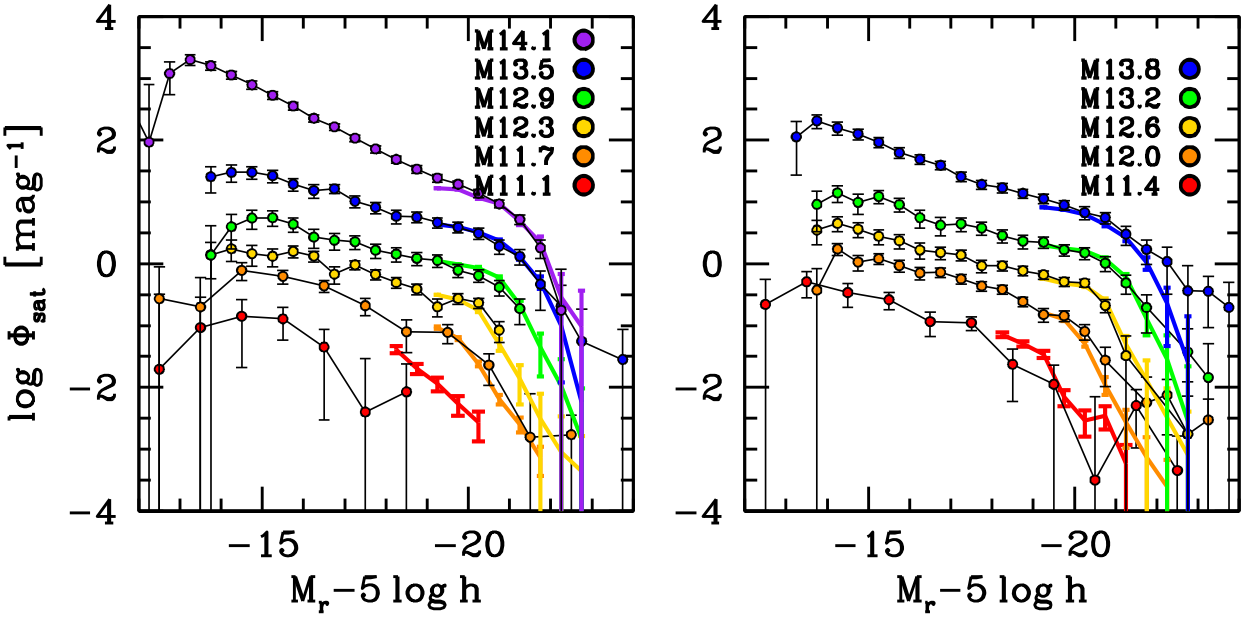

Figure 10. CLFs for satellites within spectroscopic groups. Results are split into two panels to avoid crowding. The filled circles show results using the DLIS imaging data with background subtraction. The thick coloured lines show the results using the spectroscopic redshifts obtained by matching the DLIS galaxies with SDSS spectra. Errors are from bootstrap resampling on the sample of groups. The key shows the value of $\log M_{\rm h}$. Error bars are obtained by bootstrap resampling on the set of groups.

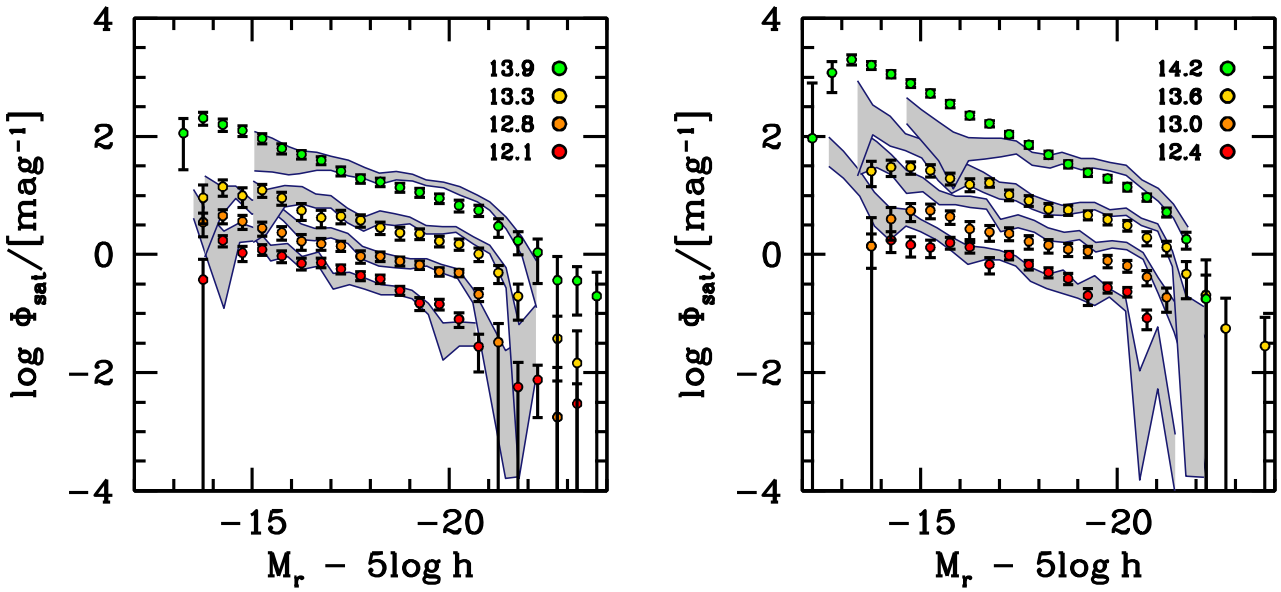

Figure 11. Same as Fig. 10, but now comparing the CLF measurements to similar measurements from Lan et al. (2016). These measurements, taken from their fig. 2, are shown with the grey shaded curves, which indicate their 1σ error ranges. The mean log halo mass for each colour is indicated in the upper-right key. The minimum halo mass of the Lan et al. results is $10^{12}\,h^{-1}$ $\mathrm{M}_{\odot}$, thus we do not compare to our lowest halo mass bins.

Previous measurements of the faint CLF in the Milky Way and M31 by Strigari & Wechsler (2012) found a cumulative number of satellites of 2.4 and 3.1, respectively, down to our magnitude limit of -14. At 10^{12} $\mathrm{M}_{\odot}$, our measurements yield a cumulative number of 2.9 satellites, which is in excellent agreement with these independent results.

In Appendix B, we present fitting functions for the measured CLFs, where the parameters of a double Schechter function are presented as functions of halo mass.

6.3 Total satellite luminosity

Fig. 13 compares the total integrated satellite luminosity, $L_{\rm sat}$, measured in the data to our abundance matching predictions. All results use our fiducial limiting magnitude of $M_r - 5\log h = -14$. The left-hand panel shows the comparison for $L^R_{\rm sat}$. The SDSS sample uses the volume-limited group catalogues, with $P_{\rm sat} < 0.1$ and halo masses estimated by the group finder. Error bars are from bootstrap resampling on the sample of groups. The dotted black line is a power-law fit to these data, which yields a scaling of $L^R_{\rm sat} \propto M_{\rm h}^{1.1}$.

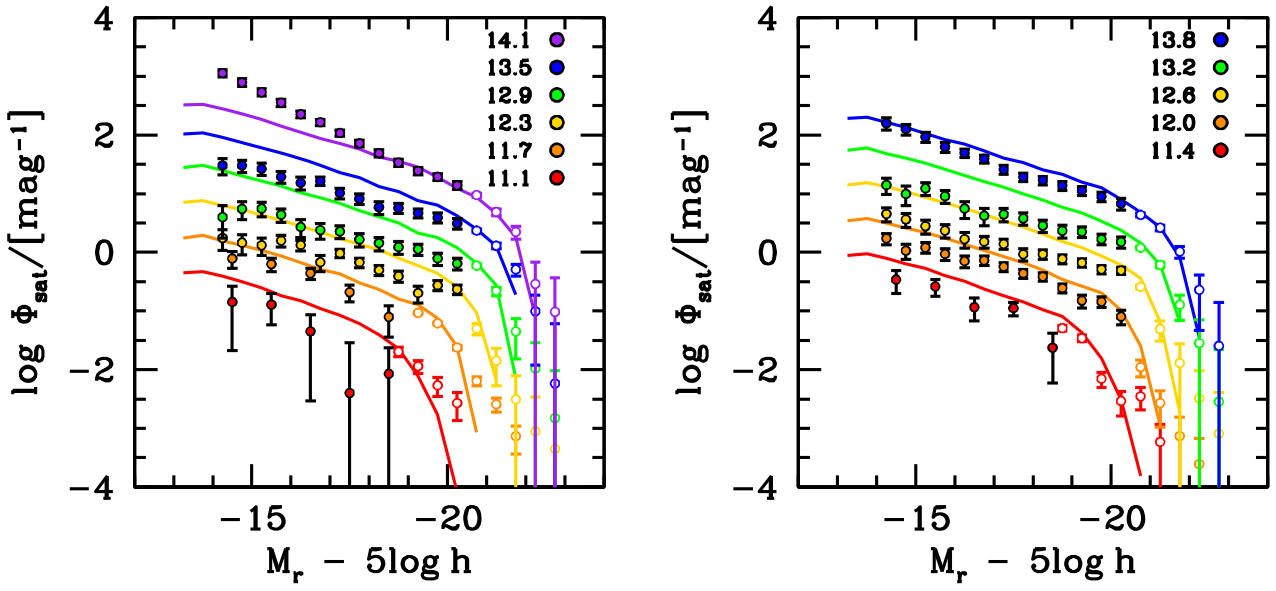

Figure 12. Same as Fig. 10, but now comparing the CLF measurements to abundance matching predictions of the C125 simulation. Solid curves show the *N*-body predictions. Circles indicate the DLIS measurements, with open circles being measurements using the DLIS galaxies with SDSS spectra, restricted to the magnitude range where the sample is complete.

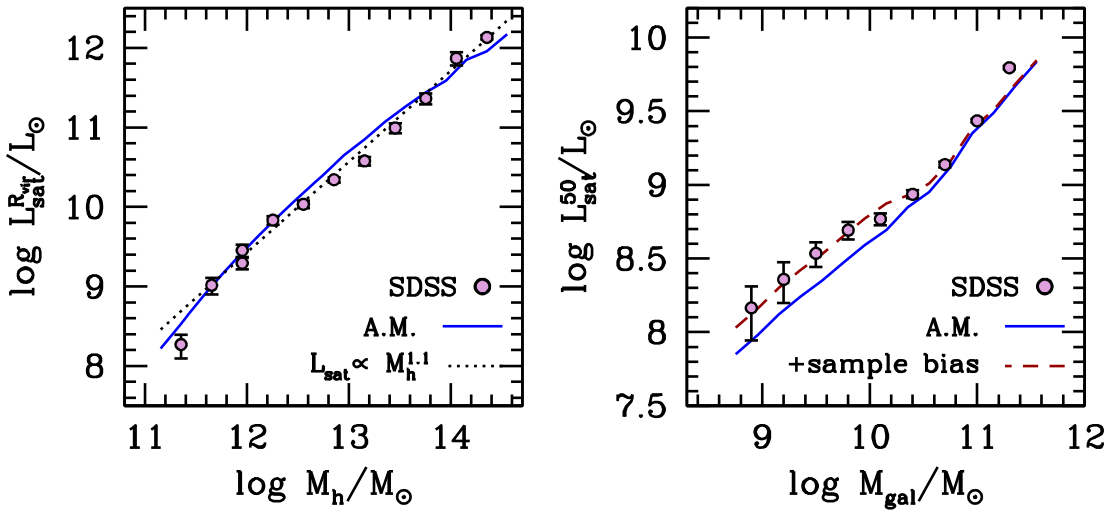

Figure 13. Left-hand panel: a comparison between $L^R_{\rm sat}$ measured around SDSS central galaxies and $L^R_{\rm sat}$ predicted in ΛCDM simulations combined with abundance matching ('A.M.'). The dotted line is the best-fitting power law to the SDSS measurements, indicating that $L^R_{\rm sat}$ scales roughly linearly with $M_{\rm h}$. Right-hand panel: same as the left-hand panel, but only now binning by central galaxy stellar mass. The dashed curve shows the prediction of abundance matching after adding the expected bias in $L^{50}_{\rm sat}$ yielded by the central galaxy finder, calibrated from our mock galaxy catalogues.

By a simple χ^2 statistic, a power law is a statistically acceptable description of these data. The solid blue curve shows the abundance matching prediction. The ΛCDM prediction scales close to a power law, being steeper at $M_{\rm h} < 10^{13}$ M$_\odot$ and shallower above this scale.

The right-hand panel of Fig. 13 compares our measurements of $L^{50}_{\rm sat}$ to the abundance matching predictions. The observational results use the full flux-limited SDSS catalogue, with central galaxies identified using the central-finding algorithm. Results are shown as a function of M_*. The blue curve once again shows our abundance matching prediction. At low stellar masses, $M_* < 10^{10.5}$ M$_\odot$, the measurements are above the predictions by about 0.2 dex. As demonstrated in Section 4, the central-finding algorithm induces some impurities in the sample of central galaxies. Using mocks, the impact of these impurities on $L^{50}_{\rm sat}$ was quantified. The dashed curve in this panel adds the bias on $L^{50}_{\rm sat}$ to the abundance matching prediction, bringing the theory and observations into near perfect agreement.

6.4 Comparison to weak lensing measurements

Fig. 14 compares our $L^{50}_{\rm sat}$ measurements to weak lensing halo mass estimates from Mandelbaum et al. (2016). The Mandelbaum et al. results split the galaxy sample by colour, with red and blue galaxies separated at $g - r = 0.8$. To facilitate a proper comparison, we convert $M_{\rm h}$ to $L^{50}_{\rm sat}$ using the abundance matching results in Fig. 1. The weak lensing results are shown with the red and blue shaded regions in the left-hand panel. Our $L^{50}_{\rm sat}$ measurements are shown

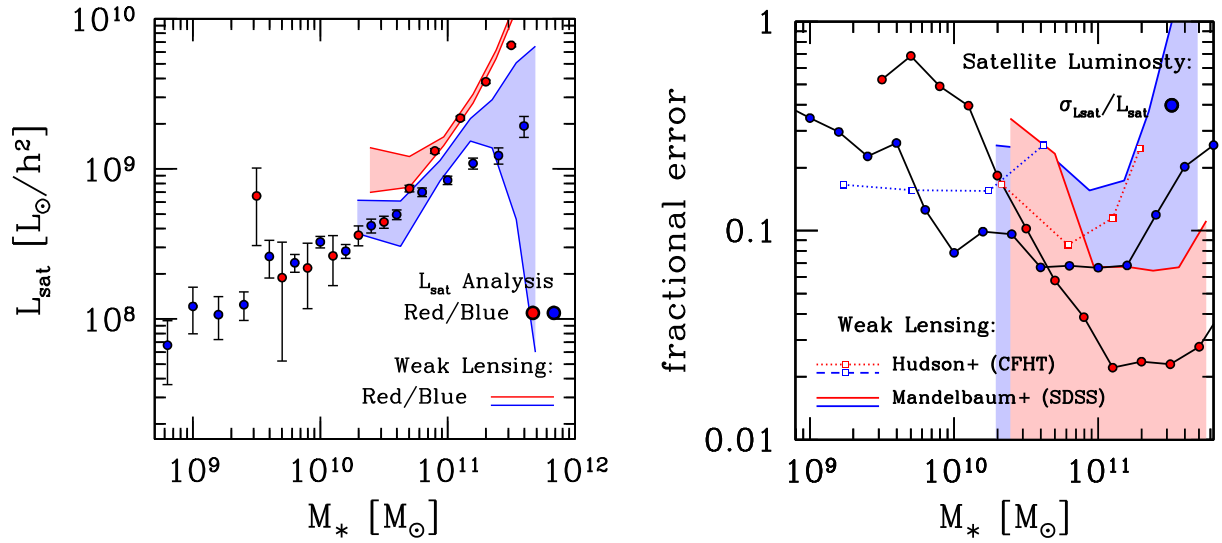

Figure 14. Left-hand panel: comparing our $L_{\rm sat}^{50}$ measurements to weak lensing measurements of SDSS central galaxies from Mandelbaum et al. (2016). Points with errors show our $L_{\rm sat}^{50}$ measurements. At each M_*, galaxies are divided by the same colour cut as used in Mandelbaum et al., $g-r>0.8$. The shaded regions indicate the weak lensing constraints. We convert the halo mass estimates of Mandelbaum et al. to $L_{\rm sat}^{50}$ values using the abundance matching results in Fig. 1. Right-hand panel: a comparison of the errors between $L_{\rm sat}^{50}$ and weak lensing. The filled circles are from our $L_{\rm sat}^{50}$ measurements, now broken into star-forming and quiescent ('Red/Blue') samples. The shaded regions indicate the errors on weak lensing halo masses by Mandelbaum et al., which is performed on the same SDSS sample as our $L_{\rm sat}^{50}$ results. The open squares connected by dotted lines show results from the Canada–France–Hawaii Telescope (CFHT) Legacy Survey by Hudson et al. (2015), a smaller area but deeper imaging.

with the points with errors. Red and blue circles indicate results for the red and blue subsamples, split with the same $g-r$ colour cut. As with the weak lensing results, the $L_{\rm sat}$ measurements indicate that red and blue central galaxies live in haloes of different mass, and this difference gets larger as M_* increases. The relative values of $L_{\rm sat}^{50}$ for the blue and red subsamples are in good agreement between the two independent approaches.

The left-hand side of Fig. 14 also highlights how $L_{\rm sat}$ is complementary to weak lensing, in that the $L_{\rm sat}$ measurements can be made to much lower values of M_*, below the limiting mass for lensing. At these lower masses, the $L_{\rm sat}^{50}$ values for the blue and red subsamples converge, indicating that low-mass central galaxies on the red sequence live in the same haloes as star-forming central galaxies of the same M_*.

The right-hand panel of Fig. 14 compares the errors on $L_{\rm sat}^{50}$ to those obtained from weak lensing mass estimates. At low M_*, the error on red galaxies for both lensing and $L_{\rm sat}^{50}$ is significantly higher due to the lower overall number of red galaxies. This reverses at the same location for both methods, $M_* \sim 4\times10^{10}$ $M_\odot$. But the errors on $L_{\rm sat}^{50}$ are roughly a factor of 5 lower than those on the weak lensing masses. We also compare to results from the Canada–France–Hawaii Telescope (CFHT) Legacy Survey from Hudson et al. (2015), representing a much smaller area survey but higher quality imaging. The result of this survey design leads to better errors than the SDSS results at low M_*, but the smaller volume limits the upper mass limit at which robust measurements can be obtained. But the errors on $L_{\rm sat}^{50}$ are still significantly lower than those from the CFHT at most masses.

7 SUMMARY

In this paper, we have presented $L_{\rm sat}$ as a proxy for dark matter haloes. The approach is to measure the total luminosity in satellite galaxies around a central galaxy. $L_{\rm sat}$ is significantly more sensitive to $M_{\rm h}$ than the luminosity or mass of the central galaxy, with a nearly linear dependence between $L_{\rm sat}^{R}$ and $M_{\rm h}$. To minimize any priors on the halo mass one would infer from this method, we also explore $L_{\rm sat}$ measured within fixed apertures of 100 and 50 h^{-1} kpc. At $M_* < 10^{10.5}$ $M_\odot$ and $M_{\rm h} < 10^{12.5}$ $M_\odot$, there is minimal impact on the scaling of $L_{\rm sat}$ when enforcing a fixed aperture. At larger masses, enforcing a fixed aperture makes $L_{\rm sat}$ less sensitive to halo mass than estimating $L_{\rm sat}$ using all satellites within the halo, but there is a still a monotonic relationship between $L_{\rm sat}$ and M_* at all stellar masses.

A number of tests and comparisons demonstrate the robustness of our approach.

(i) We find good agreement between theoretical predictions of $L_{\rm sat}$ from abundance matching models and our measurements of $L_{\rm sat}$ around SDSS galaxies.

(ii) We find good agreement between the CLFs based on our $L_{\rm sat}$ approach using photometric data and the CLF obtained from spectroscopic-only data, where the two overlap.

(iii) We find good agreement between $L_{\rm sat}$ and weak lensing results for the relative halo masses of blue and red SDSS galaxies at fixed M_*.

(iv) Given the additional signal-to-noise ratio of the $L_{\rm sat}$ method, we are able to extend this comparison of red and blue galaxies to galaxies that are more than an order of magnitude lower mass.

Additionally, at higher stellar masses, the precision of the $L_{\rm sat}$ method allows us to compare galaxies not just in divisions of red and blue, but then to further fine bin these divisions based on other galaxy properties. Alpaslan & Tinker (2020), a companion paper to this paper, present these initial results from SDSS.

The main theoretical uncertainty in using $L_{\rm sat}$ as a halo mass proxy is the degeneracy of $L_{\rm sat}$ on halo formation history; late-forming haloes have more substructure and thus larger amounts of satellite luminosity. However, this correlation makes a distinct prediction for the clustering and large-scale environments of central galaxies. For example, Tinker et al. (2018b), looking at galaxies on the star-forming

main sequence, showed that star formation rate correlated with large-scale environment at fixed stellar mass: above-average star-forming galaxies live in below-average large-scale densities. These results were consistent with a model in which galaxy star formation rate was correlated with dark matter halo accretion rate. Thus, if $L_{\rm sat}$ correlates with star formation rate, the measurements of Tinker et al. (2018b) can break the degeneracy, and quantify how much of the $L_{\rm sat}$ variation is due to a change in $M_{\rm h}$, how much is due to a correlation with $z_{1/2}$. But this is only for star-forming galaxies. In contrast, when dividing galaxies into star-forming and quiescent samples, the quenched fraction of central galaxies are *independent* of environment (Tinker et al. 2008, 2017, 2018a; Peng et al. 2010; Zu & Mandelbaum 2016, 2018; Wang et al. 2018). Thus, if there are differences in $L_{\rm sat}$ between red and blue galaxies, this is truly due to differences in $M_{\rm h}$, and not correlated with $z_{1/2}$. This approach can be extended to any galaxy property one wishes to probe, including galaxy size, velocity dispersion, and morphology.

The primary observational systematic is in the definition of the sample of central galaxies. Impurities in this sample always go in the direction of increasing $L_{\rm sat}$ at fixed M_*. No method of identifying central galaxies will be perfect, but the two methods used here have minimal impact on $L_{\rm sat}$, increasing it by 0.1–0.2 dex. The primary concern is when measuring relative values of $L_{\rm sat}$ when splitting up galaxies by secondary properties at fixed M_*. If the secondary property correlates with satellite fraction, the impact of impurities will have a differential effect on the relative values of $L_{\rm sat}$. The division that should maximize this error – splitting the sample based on colour or star formation bimodality – still only produced a bias of 0.1–0.2 dex. The results from our mock galaxy tests can also be used to correct observational results, or set systematic error bars.

The $L_{\rm sat}$ technique opens a new window into the galaxy–halo connection. Many of the outstanding issues in this relationship can be addressed with $L_{\rm sat}$. In a recent review of the galaxy–halo connection, Wechsler & Tinker (2018) highlight the need for additional data to constrain the relationship between halo mass and secondary galaxy properties, including galaxy colour and galaxy size. However, these secondary relationships are difficult to detect through previous methods. Given these open questions, the $L_{\rm sat}$ technique has many applications that are complementary with direct probes of halo mass.

ACKNOWLEDGEMENTS

The simulations used in this study were produced with computational resources at SLAC National Accelerator Laboratory, a U.S. Department of Energy Office; Y-YM and RHW thank Matthew R. Becker for creating these simulations, and thank the support of the SLAC computational team.

Y-YM was supported by the Samuel P. Langley PITT PACC Postdoctoral Fellowship, and by NASA through the NASA Hubble Fellowship grant no. *HST*-HF2-51441.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555.

This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02- 05CH11231.

The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; NOAO Proposal ID # 2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing–Arizona Sky Survey (BASS; NOAO Proposal ID # 2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; NOAO Proposal ID # 2016A-0453; PI: Arjun Dey). DECaLS, BASS, and MzLS together include data obtained, respectively, at the Blanco Telescope, Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory (NOAO); the Bok Telescope, Steward Observatory, University of Arizona; and the Mayall Telescope, Kitt Peak National Observatory, NOAO. The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

NOAO is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute for Cosmological Physics at the University of Chicago, Center for Cosmology and Astroparticle Physics, Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnológico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft, and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenossische Technische Hochschule (ETH) Zurich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universitat Munchen and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories, Chinese Academy of Sciences (the Strategic Priority Research Program 'The Emergence of Cosmological Structures' Grant # XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. The BASS is also supported by the External Cooperation Program of Chinese Academy of Sciences (Grant # 114A11KYSB20160057), and National Natural Science Foundation of China (Grant # 11433005).

The Legacy Survey team makes use of data products from the *Near-Earth Object Wide-field Infrared Survey Explorer* (*NEOWISE*), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. *NEOWISE* is funded by the National Aeronautics and Space Administration.

The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics

of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123, by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to NOAO.

DATA AVAILABILITY

All data from the DESI Legacy Imaging Surveys (DLIS) are available through their web page. All SDSS data are available through the public SDSS releases, with specific URLs listed in this paper. $L_{\rm sat}$ measurements are available upon request of the authors.

REFERENCES

Abbas U., Sheth R. K., 2006, MNRAS, 372, 1749
Alpaslan M., Tinker J. L., 2020, MNRAS, 496, 5463
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, ApJ, 762, 109
Behroozi P. S., Wechsler R. H., Conroy C., 2013b, ApJ, 762, L31
Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013c, ApJ, 763, 18
Behroozi P. S., Wechsler R. H., Conroy C., 2013d, ApJ, 770, 57
Blanton M. R. et al., 2005a, AJ, 129, 2562
Blanton M. R., Lupton R. H., Schlegel D. J., Strauss M. A., Brinkmann J., Fukugita M., Loveday J., 2005b, ApJ, 631, 208
Campbell D., van den Bosch F. C., Hearin A., Padmanabhan N., Berlind A., Mo H. J., Tinker J., Yang X., 2015, MNRAS, 452, 444
Cao J.-z., Tinker J. L., Mao Y.-Y., Wechsler R. H., 2020, MNRAS, 498, 5080
Chen Y.-M. et al., 2012, MNRAS, 421, 314
Coil A. L. et al., 2011, ApJ, 741, 8
Conroy C. et al., 2005, ApJ, 635, 990
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
DeRose J. et al., 2019, preprint (arXiv:1901.02401)
Dey A. et al., 2019, AJ, 157, 168
Flaugher B. et al., 2015, AJ, 150, 150
Gao L., White S. D. M., 2007, MNRAS, 377, L5
Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004, MNRAS, 355, 819
Geha M. et al., 2017, ApJ, 847, 4
Giocoli C., Tormen G., Sheth R. K., 2012, MNRAS, 422, 185
Grogin N. A. et al., 2011, ApJS, 197, 35
Han J., Cole S., Frenk C. S., Jing Y., 2016, MNRAS, 457, 1208
Han J., Li Y., Jing Y., Nishimichi T., Wang W., Jiang C., 2019, MNRAS, 482, 1900
Hansen S. M., Sheldon E. S., Wechsler R. H., Koester B. P., 2009, ApJ, 699, 1333
Hearin A. P., Watson D. F., 2013, MNRAS, 435, 1313
Hudson M. J. et al., 2015, MNRAS, 447, 298
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, MNRAS, 395, 160
Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
Lan T.-W., Ménard B., Mo H., 2016, MNRAS, 459, 3998
Lange J. U., van den Bosch F. C., Zentner A. R., Wang K., Villarreal A. S., 2019, MNRAS, 482, 4824
Leauthaud A. et al., 2012, ApJ, 744, 159
Lehmann B. V., Mao Y.-Y., Becker M. R., Skillman S. W., Wechsler R. H., 2017, ApJ, 834, 37
Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, MNRAS, 391, 1940
MacCrann N. et al., 2018, MNRAS, 480, 4614
Man Z.-Y., Peng Y.-J., Shi J.-J., Kong X., Zhang C.-P., Dou J., Guo K.-X., 2019, ApJ, 881, 74
Mandelbaum R., Wang W., Zu Y., White S., Henriques B., More S., 2016, MNRAS, 457, 3200
Mansfield P., Kravtsov A. V., 2020, MNRAS, 493, 4763
Mao Y.-Y., Williamson M., Wechsler R. H., 2015, ApJ, 810, 21
Mao Y.-Y., Zentner A. R., Wechsler R. H., 2018, MNRAS, 474, 5143
Masjedi M. et al., 2006, ApJ, 644, 54
Matthee J., Schaye J., Crain R. A., Schaller M., Bower R., Theuns T., 2017, MNRAS, 465, 2381
More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, MNRAS, 410, 210
Nadler E. O., Mao Y.-Y., Green G. M., Wechsler R. H., 2019, ApJ, 873, 34
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Norberg P., Frenk C. S., Cole S., 2008, MNRAS, 383, 646
Peng Y.-j. et al., 2010, ApJ, 721, 193
Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, 771, 30
Rodríguez-Puebla A., Avila-Reese V., Yang X., Foucaud S., Drory N., Jing Y. P., 2015, ApJ, 799, 130
Salcedo A. N., Maller A. H., Berlind A. A., Sinha M., McBride C. K., Behroozi P. S., Wechsler R. H., Weinberg D. H., 2018, MNRAS, 475, 4411
Sales L. V., Wang W., White S. D. M., Navarro J. F., 2013, MNRAS, 428, 573
Sinha M., Berlind A. A., McBride C. K., Scoccimarro R., Piscionere J. A., Wibking B. D., 2018, MNRAS, 478, 1042
Springel V. et al., 2008, MNRAS, 391, 1685
Strauss M. A. et al., 2002, AJ, 124, 1810
Strigari L. E., Wechsler R. H., 2012, ApJ, 749, 75
Tal T., Wake D. A., van Dokkum P. G., van den Bosch F. C., Schneider D. P., Brinkmann J., Weaver B. A., 2012, ApJ, 746, 138
Tinker J. L., Norberg P., Weinberg D. H., Warren M. S., 2007, ApJ, 659, 877
Tinker J. L., Conroy C., Norberg P., Patiri S. G., Weinberg D. H., Warren M. S., 2008, ApJ, 686, 53
Tinker J., Wetzel A., Conroy C., 2011, preprint (arXiv:1107.5046)
Tinker J. L. et al., 2012, ApJ, 745, 16
Tinker J. L., Leauthaud A., Bundy K., George M. R., Behroozi P., Massey R., Rhodes J., Wechsler R. H., 2013, ApJ, 778, 93
Tinker J. L., Wetzel A. R., Conroy C., Mao Y.-Y., 2017, MNRAS, 472, 2504
Tinker J. L., Hahn C., Mao Y.-Y., Wetzel A. R., Conroy C., 2018a, MNRAS, 477, 935
Tinker J. L., Hahn C., Mao Y.-Y., Wetzel A. R., 2018b, MNRAS, 478, 4487
Tojeiro R. et al., 2017, MNRAS, 470, 3720
Vakili M., Hahn C., 2019, ApJ, 872, 115
Walsh K., Tinker J., 2019, MNRAS, 488, 470
Wang W., White S. D. M., 2012, MNRAS, 424, 2574
Wang H. et al., 2018, ApJ, 852, 31
Wang K. et al., 2019, MNRAS, 488, 3541
Watson D. F., Berlind A. A., McBride C. K., Hogg D. W., Jiang T., 2012, ApJ, 749, 83
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52
Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, ApJ, 652, 71
Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, MNRAS, 366, 2
Wetzel A. R., Tinker J. L., Conroy C., 2012, MNRAS, 424, 232
Yang X., Mo H. J., van den Bosch F. C., Jing Y. P., 2005, MNRAS, 356, 1293
Yang X., Mo H. J., van den Bosch F. C., 2008, ApJ, 676, 248
Zehavi I. et al., 2005, ApJ, 630, 1
Zehavi I. et al., 2011, ApJ, 736, 59
Zentner A. R., Berlind A., Bullock J. S., Kravtsov A., Wechsler R. H., 2005, ApJ, 624, 505
Zentner A. R., Hearin A., van den Bosch F. C., Lange J. U., Villarreal A., 2019, MNRAS, 485, 1196
Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760
Zhou Z. et al., 2018, PASP, 130, 085001
Zu Y., Mandelbaum R., 2015, MNRAS, 454, 1161
Zu Y., Mandelbaum R., 2016, MNRAS, 457, 4360
Zu Y., Mandelbaum R., 2018, MNRAS, 476, 1637

APPENDIX A: FINDING CENTRAL GALAXIES

Here we present the details of our central-finding algorithm. The purpose of this algorithm is to determine if a galaxy is likely to be a central, but without attempting to determine the mass of its host halo. The basic approach is very similar to running to the initialization stage of our group finder, which works on inverse abundance matching – i.e. rather than putting mock galaxies within N-body dark matter haloes, we put haloes around observed galaxies by the same abundance matching calculation:

$$\int_{M_*}^{\infty} \Phi(M_*{}')\,\mathrm{d}M_*{}' = \int_{M_{\rm h}}^{\infty} n(M_{\rm h}')\,\mathrm{d}M_{\rm h}', \tag{A1}$$

where $\Phi(M_*)$ is the observed stellar mass function and $n(M_{\rm h})$ is the mass function of dark matter haloes. Equation (A1) is the simplest form of abundance matching, assuming no scatter between M_* and $M_{\rm h}$. The first step in our group finder is to use equation (A1) as a first guess of the haloes around each galaxy, regardless of whether they are centrals or satellites, then begin the process of determining the probability that each galaxy is a satellite within a larger dark matter haloes. In the group finder, the halo mass is itself abundance matched on to the *total group stellar mass*, and the entire sample is iterated to convergence.

In our central finder, we make the approach even more simple and flexible by using pre-tabulated abundance matching relations between $M_{\rm h}$ and M_*. In our fiducial approach, we use the relations tabulated in Behroozi et al. (2013d), which quantify the SHMR from $z = 0$ to $z = 8$. The use of pre-tabulated relations means that it is no longer necessary to perform the abundance matching on volume-limited samples of galaxies; a 10^{11} $\mathrm{M}_\odot$ central galaxy at $z = 0.02$, which likely has tens of satellites within the flux-limited SDSS MGS, is assigned the same halo mass at the same galaxy observed at $z = 0.15$, where it likely has no satellite galaxies.

Just as in the group finder, the probability that a galaxy is a satellite in a larger dark matter halo is given by

$$P_{\rm cen} = \left(1 + P_{R_{\rm p}} P_{\Delta z}/B\right)^{-1}, \tag{A2}$$

where $P_{R_{\rm p}}$ is the probability at a given projected separation from the centre of the halo, and $P_{\Delta z}$ is the probability at a given line-of-sight separation from the redshift of the halo. B is a constant determined from calibration on mock galaxy samples, set to be $B = 10$. The former is given by the projected NFW density profile (Navarro et al. 1997), while the latter assumes a Gaussian probability distribution function with width given by the virial velocity dispersion of the host halo. Further details can be found in appendix A of Tinker et al. (2011) or Yang et al. (2005), on which our algorithm is based.

The drawback to using the inverse abundance-matching approach is that the haloes assigned to galaxies will be biased, given the fact that asymmetric scatter yields different mean relationships of $\langle M_*|M_{\rm h}\rangle$ and $\langle M_{\rm h}|M_*\rangle$. However, this error generally accrues to overestimate the masses of host haloes, thus making it a conservative approach to determining if nearby galaxies are satellites. The other drawback to this approach is that stellar mass estimates can differ widely, thus the estimate used in Behroozi et al. (2013d) may differ from that used in a given sample of galaxies. However, in our tests we find that this yields minimal bias.

Fig. A1 shows the results of our central finder when applied to four different galaxy mocks. As a baseline, the left-hand panel shows results when applied to a volume-limited sample of $z = 0$ galaxies, complete down to a stellar mass of $M_* = 10^{9.3}$ $\mathrm{M}_\odot$. The stellar mass function used to put stellar mass into the haloes is the PCA stellar mass function. Normally, one would use the full group finder on a volume-limited sample, but this test provides a good comparison for how results degrade in less optimal survey samples. The second panel is a flux-limited sample with same redshift distribution as the SDSS MGS. For $P_{\rm sat} < 0.3$, the sample of central galaxies identified

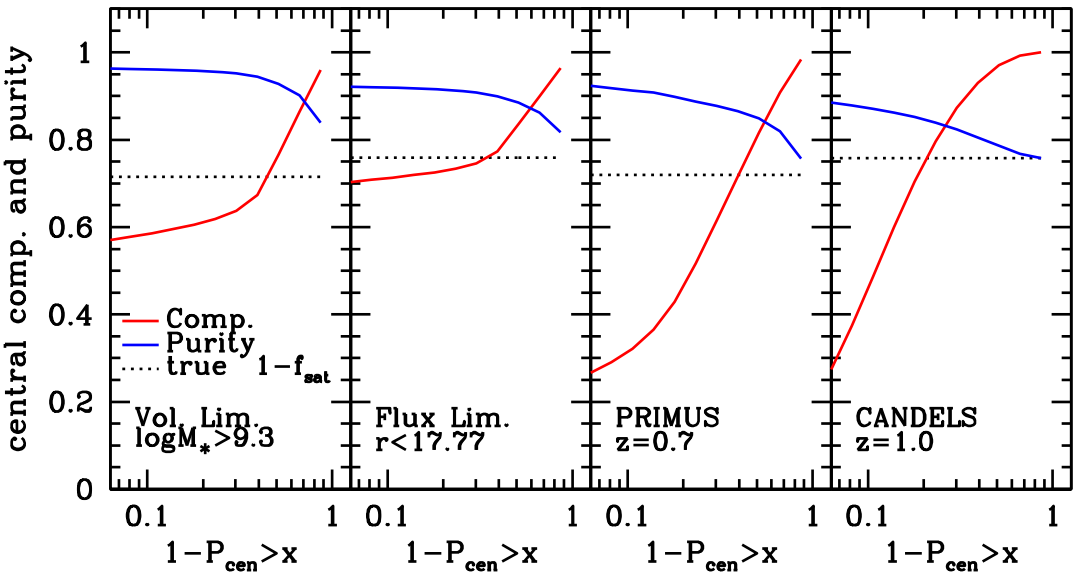

Figure A1. The purity and completeness of central galaxies identified by our central-finding algorithm. The solid curves show the results as a function of $P_{\rm sat}$ threshold. The dotted line shows the true fraction of central galaxies in each mock galaxy catalogue. From left to right, the panels are: a volume-limited sample of galaxies, complete to a stellar mass of $M_* = 10^{9.3}$ $\mathrm{M}_\odot$. A flux-limited sample of galaxies with the same flux limit as the SDSS MGS, and the same $n(z)$. A stellar-mass complete sample of galaxies at $z = 0.7$ with redshift errors comparable to those in the PRIsm MUlti-object Survey (PRIMUS). A stellar-mass complete sample of galaxies at $z = 1.0$ with photometric redshift errors set to equal those from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS).

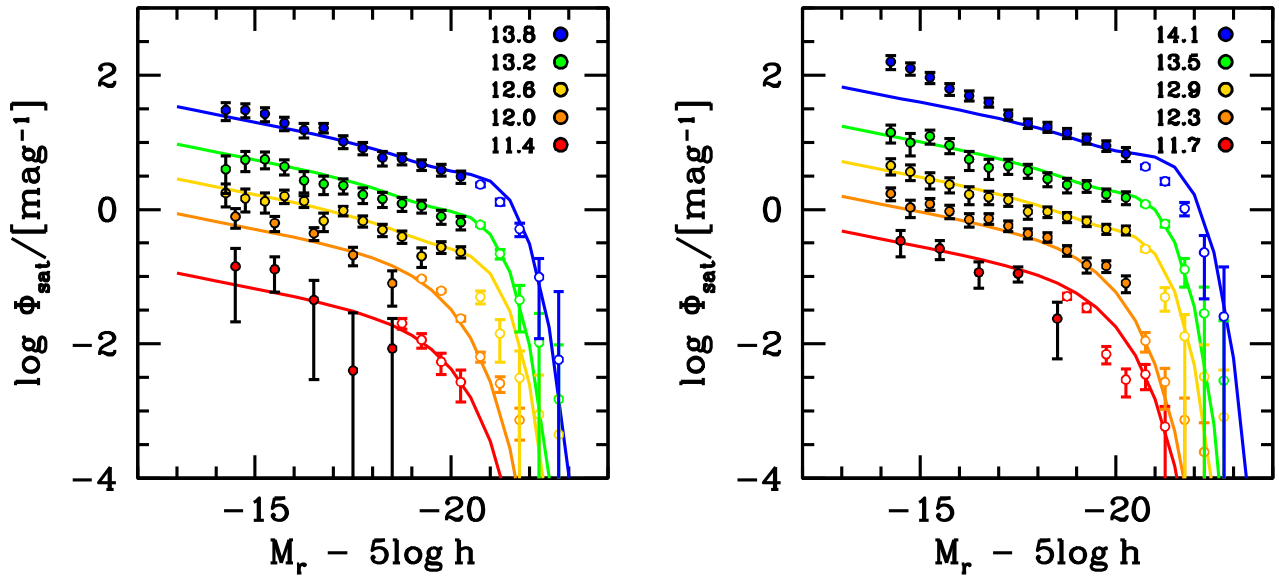

Figure B1. Same as Fig. 10, but now comparing the CLF measurements to double Schechter function fits. The parameters of the fitting function vary with $\log M_{\rm h}/{\rm M}_\odot$. The fits are optimized for halo mass bins $\log M_{\rm h} < 14.1$. Larger haloes suffer from insufficient background subtraction at the very faint end of the CLF.

by the algorithm has a purity of > 90 per cent and a completeness of around 70 per cent.

The two right-hand panels show the algorithm as applied to mock samples that have significantly degraded redshift information. Here, we replace the host halo velocity dispersion used the halo in $P_{\Delta z}$ with the average error on the redshift. To construct mock samples we use the Buzzard mocks of DeRose et al. (2019). The rightmost panel is a mock sample comparable to the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) at $z = 1.0$ (Grogin et al. 2011); it is complete down to a stellar mass of $M_* = 10^{9.5}$ with a photometric redshift error of $\sigma_z/(1+z) = 0.033$. The next rightmost panel is an approximation of the redshift accuracy and redshift range of the PRIsm MUlti-object Survey (PRIMUS; Coil et al. 2011), with $z \approx 0.7$ and $\sigma_z/(1+z) = 0.005$. For both of these mock surveys, the algorithm yields central galaxy samples that are complete to ∼90 per cent at $P_{\rm sat} < 0.1$. However, this comes at a significant cost of the completeness of the sample, which dips below 50 per cent at this $P_{\rm sat}$ threshold.

This code and a short instruction manual is made publicly available.[4]

APPENDIX B: FITTING THE CONDITIONAL LUMINOSITY FUNCTIONS

We model the conditional luminosity function (CLF) of satellites within haloes using the modified Schechter function employed by Blanton et al. (2005b) to describe the luminosity function of low-luminosity galaxies in the SDSS. This fitting function takes the form

$$\Phi(M) = 0.4 \ln 10 \, {\rm d}M \exp\left(-10^{-0.4(M-M_*)}\right) \times \left[\phi_{*,1} 10^{-0.4(M-M_*)(\alpha_1+1)} + \phi_{*,2} 10^{-0.4(M-M_*)(\alpha_2+1)}\right]. \quad \text{(B1)}$$

The motivation for this function in Blanton et al. (2005b) is to better model an upturn in the luminosity function at magnitudes fainter than $M_r = -18$. A similar upturn is seen in the CLFs in higher mass haloes in our results in Fig. 10. We fix the values of the power-law indices to be of $\alpha_1 = 1$ and $\alpha_2 = -1.28$, with the latter parameter fixing the faint-end slope to be the same for all haloes. We construct fitting functions for how the other parameters of the modified Schechter function depend on halo mass:

$$\phi_{*,1} = \begin{cases} 0.98x - 12.85 & \text{if } x > 12.5, \\ 0 & \text{if } x \leq 12.5, \end{cases} \quad \text{(B2)}$$

where $x \equiv \log M_{\rm h}$ with the halo mass in units of ${\rm M}_\odot$,

$$\phi_{*,2} = \begin{cases} 0.86x - 11.10 & \text{if } x > 11.7, \\ 2.13x - 26 & \text{if } x \leq 11.7, \end{cases} \quad \text{(B3)}$$

and

$$M_* = \begin{cases} -0.99x - 6.36 & \text{if } x > 13.3, \\ -19.58 & \text{if } x \leq 13.3. \end{cases} \quad \text{(B4)}$$

Equation (B2) implies that, for haloes less massive than $10^{12.5}$ ${\rm M}_\odot$, a single power-law Schechter function is sufficient for modelling the CLF. The results of the fitting functions are shown in Fig. B1. The fitting parameters themselves were obtained for the results at $M_{\rm h} \leq 10^{13.8}$ ${\rm M}_\odot$. At higher halo masses, the faint end of the CLF deviates significantly from the trends seen in lower mass haloes. This is likely due to insufficient background subtraction for the most massive haloes, and this effect is also seen in our tests with mock galaxy samples (cf. Fig. 6).

[4] https://github.com/jltinker/IsolationCriterion