

# The 6dF Galaxy Survey: Dependence of halo occupation on stellar mass

Florian Beutler<sup>1,2\*</sup>, Chris Blake<sup>3</sup>, Matthew Colless<sup>4</sup>, D. Heath Jones<sup>5</sup>,  
Lister Staveley-Smith<sup>2,6</sup>, Lachlan Campbell<sup>7</sup>, Quentin Parker<sup>4,8</sup>, Will Saunders<sup>4</sup>,  
Fred Watson<sup>4</sup>

<sup>1</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

<sup>2</sup>*International Centre for Radio Astronomy Research (ICRAR), University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Australia*

<sup>3</sup>*Centre for Astrophysics & Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn, VIC 3122, Australia*

<sup>4</sup>*Australian Astronomical Observatory, PO Box 296, Epping NSW 1710, Australia*

<sup>5</sup>*School of Physics, Monash University, Clayton, VIC 3800, Australia*

<sup>6</sup>*ARC Centre of Excellence for All-sky Astrophysics (CAASTRO)*

<sup>7</sup>*Western Kentucky University, Bowling Green, KY 42101, USA*

<sup>8</sup>*Department of Physics and Astronomy, Faculty of Sciences, Macquarie University, NSW 2109, Sydney, Australia*

18 December 2012

## ABSTRACT

In this paper we study the stellar-mass dependence of galaxy clustering in the 6dF Galaxy Survey. The near-infrared selection of 6dFGS allows more reliable stellar mass estimates compared to optical bands used in other galaxy surveys. Using the Halo Occupation Distribution (HOD) model, we investigate the trend of dark matter halo mass and satellite fraction with stellar mass by measuring the projected correlation function,  $w_p(r_p)$ . We find that the typical halo mass ( $M_1$ ) as well as the satellite power law index ( $\alpha$ ) increase with stellar mass. This indicates, (1) that galaxies with higher stellar mass sit in more massive dark matter halos and (2) that these more massive dark matter halos accumulate satellites faster with growing mass compared to halos occupied by low stellar mass galaxies. Furthermore we find a relation between  $M_1$  and the minimum dark matter halo mass ( $M_{\min}$ ) of  $M_1 \approx 22 M_{\min}$ , in agreement with similar findings for SDSS galaxies. The satellite fraction of 6dFGS galaxies declines with increasing stellar mass from 21% at  $M_{\text{stellar}} = 2.6 \times 10^{10} h^{-2} M_{\odot}$  to 12% at  $M_{\text{stellar}} = 5.4 \times 10^{10} h^{-2} M_{\odot}$  indicating that high stellar mass galaxies are more likely to be central galaxies. We compare our results to two different semi-analytic models derived from the Millennium Simulation, finding some disagreement. Our results can be used for placing new constraints on semi-analytic models in the future, particularly the behaviour of luminous red satellites. Finally we compare our results to studies of halo occupation using galaxy-galaxy weak lensing. We find good overall agreement, representing a valuable crosscheck for these two different tools of studying the matter distribution in the Universe.

**Key words:** galaxy formation, large-scale structure of Universe, surveys, galaxies: statistics, galaxies: halos

## 1 INTRODUCTION

The first statistical studies of galaxy clustering (Totsuji & Kihara 1969; Peebles 1973; Hauser & Peebles 1973, 1974; Peebles 1974) found that the galaxy correlation function behaves like a power law, which is difficult to explain from first principles (Berlind & Weinberg 2001).

More recent studies, however, found deviations from a power law. For example Zehavi et al. (2005a) showed that the projected correlation function  $w_p(r_p)$  of SDSS galaxies exhibits a statistically significant departure from a power law. They also showed that a 3-parameter Halo Occupation Distribution (HOD) model (e.g., Jing, Mo & Borner 1998; Ma & Fry 2000; Peacock & Smith 2000; Seljak 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2001; Cooray & Sheth 2002) together with a  $\Lambda$ CDM background

\* E-mail: fbeutler@lbl.gov

cosmology, can account for this departure, reproducing the observed  $w_p(r_p)$ .

Within the halo model the transition from the 1-halo term to the 2-halo term causes a "dip" in the correlation function at around  $1 - 3h^{-1}$  Mpc, corresponding to the exponential cutoff in the halo mass function. In case of a smooth transition between the one- and two-halo terms, this can mimic a power-law correlation function. Studies with Luminous Red Galaxies (LRGs) found that the deviation from a power-law is larger for highly clustered bright galaxies (Zehavi et al. 2005a,b; Blake, Collister & Lahav 2008; Zheng et al. 2009; Zehavi et al. 2010), and at high redshift (Conroy, Wechsler & Kravtsov 2006), which agrees with theoretical predictions (Watson et al. 2011).

While galaxy clustering is difficult to predict, dark matter clustering is dominated by gravity and can be predicted for a given cosmology using N-body simulations. Using models for how galaxies populate dark matter halos, which are usually motivated by N-body simulations, we can directly link galaxy clustering and matter clustering. This can be modelled in terms of the probability distribution  $p(N|M)$  that a halo of virial mass  $M$  contains  $N$  galaxies of a given type. On strongly non-linear scales the dark matter distribution is given by the actual density distribution of the virialized halos, while on large and close to linear scales the dark matter distribution can be predicted from linear perturbation theory.

HOD modelling has been applied to galaxy clustering data from the 2-degree Field Galaxy Redshift Survey (2dFGRS) (Porciani, Magliocchetti & Norberg 2007; Tinker et al. 2006) and the Sloan Digital Sky Survey (SDSS) (van den Bosch, Yang & Mo 2003; Magliocchetti & Porciani 2003; Zehavi et al. 2005a,b; Tinker et al. 2005; Yang et al. 2005, 2008; Zehavi et al. 2010). More recently it also became possible to model the clustering of high- $z$  galaxies using VVDS (Abbas et al. 2010), Boötes (Brown et al. 2008), DEEP2 (Coil et al. 2006) and Lyman-break galaxies at high redshift in the GOODS survey (Lee et al. 2006). Such studies revealed, that the minimum mass,  $M_{\min}$  for a halo to host a central galaxy more luminous than some threshold,  $L$  is proportional to  $L$  at low luminosities, but steepens above  $L_*$ . Massive halos have red central galaxies with predominantly red satellites, while the fraction of blue central galaxies increases with decreasing host halo mass. Furthermore Zehavi et al. (2005b) found that there is a scaling relation between the minimum mass of the host halos,  $M_{\min}$  and the mass scale,  $M_1$  of halos that on average host one satellite galaxy in addition to the central galaxy,  $M_1 \approx 23 M_{\min}$ . Using a different HOD parameterization, Zheng, Coil & Zehavi (2007) found the relation to be  $M_1 \approx 18 M_{\min}$ , very similar to Zehavi et al. (2010) who found  $M_1 \approx 17 M_{\min}$ .

The 6dF Galaxy Survey is one of the biggest galaxy surveys available today with a sky coverage of 42% and an average redshift of  $\bar{z} = 0.05$ . The survey includes about 125 000 redshifts selected in the  $J, H, K, b_J, r_F$ -bands (Jones et al. 2004, 2005, 2009). The near infrared selection, the high completeness and the wide sky coverage make 6dFGS one of the best surveys in the local Universe to study galaxy formation. This dataset has been used to study the large scale galaxy clustering to measure the Hubble constant using Baryon Acoustic Oscillations (Beutler et al. 2011), as

well as the growth of structure at low redshift (Beutler et al. 2012). While these previous studies used the  $K$ -band selected sample, in this analysis we use the  $J$ -band. The  $J$ -band allows the most reliable stellar mass estimate of the five bands available in 6dFGS, because of its lower background noise. Together with the  $b_J - r_F$  colour we can derive stellar masses using the technique of Bell & De Jong (2001), which leads to a dataset of 76 833 galaxies in total. The photometric near-infrared selection from 2MASS makes the stellar mass estimates in 6dFGS more reliable than stellar mass estimates in other large galaxy surveys which rely on optical bands (Drory et al. 2004; Kannappan & Gawiser 2007; Longhetti & Saracco 2009; Grillo et al. 2008; Gallazzi & Bell 2009).

Numerical N-body simulations are usually restricted to dark matter only. To understand galaxy formation, baryonic effects such as feedback and gas cooling, have to be included. Such simulations face severe theoretical and numerical challenges. Semi-analytic models build upon pre-calculated dark matter merger trees from cosmological simulations and include simplified, physically and observationally motivated, analytic recipes for different baryonic effects. Semi-analytic models have been shown to successfully reproduce observed statistical properties of galaxies over a large range of galaxy masses and redshifts (e.g. Croton et al. 2006; Bower et al. 2006; De Lucia & Blaizot 2007; Bertone et al. 2007; Font et al. 2008; Guo et al. 2010) and allow a level of understanding unavailable in N-body simulations. The underlying models are necessarily simplified and often use a large number of free parameters to fit different observations simultaneously. In this paper, we derive 6dFGS mock surveys from semi-analytic models based on the Millennium Simulation (Springel et al. 2005) and compare the properties of these surveys with measurements in 6dFGS. Our results can be used to improve upon these semi-analytic models and further our understanding of baryonic feedback processes on galaxy clustering.

Mandelbaum et al. (2006) studied halo occupation as a function of stellar mass and galaxy type using galaxy-galaxy weak lensing in SDSS. They found that for a given stellar mass, the halo mass is independent of morphology below  $M_{\text{stellar}} = 10^{11} M_{\odot}$ , indicating that stellar mass is a good proxy for halo mass at that range. We compare our results with Mandelbaum et al. (2006) which represents a valuable crosscheck of the HOD analysis using two very different techniques, weak lensing and galaxy clustering.

This paper is organised as follows: First we introduce the 6dF Galaxy Survey in section 2 together with the technique to derive the stellar masses. We also explain how we derive our four volume-limited sub-samples in stellar mass and redshift, which are then used for the further analysis. In section 3 we calculate the projected correlation function,  $w_p(r_p)$ , for each sub-sample and use jack-knife re-sampling to derive the covariance matrices. In section 4 we fit power laws to the projected correlation functions of the four sub-samples. In section 5 we introduce the HOD framework and in section 6 we apply the HOD model to the data. In section 7 we derive 6dFGS mock samples from two different semi-analytic models, which we then compare to our results in section 8 together with a general discussion of our findings. We conclude in section 9.

Throughout the paper we use  $r$  to denote real space sep-

arations and  $s$  to denote separations in redshift space. Our fiducial model to convert redshifts into distances is a flat universe with  $\Omega_m = 0.27$ ,  $w = -1$  and  $\Omega_k = 0$ . The Hubble constant is set to  $H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$  which sets the unit of stellar masses to  $h^{-2} M_\odot$ , while most other masses are given in  $h^{-1} M_\odot$ . The HOD model uses cosmological parameters following WMAP7 (Komatsu et al. 2011).

## 2 THE 6DF GALAXY SURVEY

The 6dF Galaxy Survey (6dFGS; Jones et al. 2004, 2005, 2009) is a near-infrared selected ( $J, H, K$ ) redshift survey covering  $17\,000 \text{ deg}^2$  of the southern sky. The  $J$ ,  $H$  and  $K$  surveys avoids a  $\pm 10^\circ$  region around the Galactic Plane to minimise Galactic extinction and foreground source confusion in the Plane. The near-infrared photometric selection was based on total magnitudes from the Two-Micron All-Sky Survey Extended Source Catalog (2MASS XSC; Jarrett et al. 2000). The spectroscopic redshifts of 6dFGS were obtained with the Six-Degree Field (6dF) multi-object spectrograph of the UK Schmidt Telescope (UKST) between 2001 and 2006.

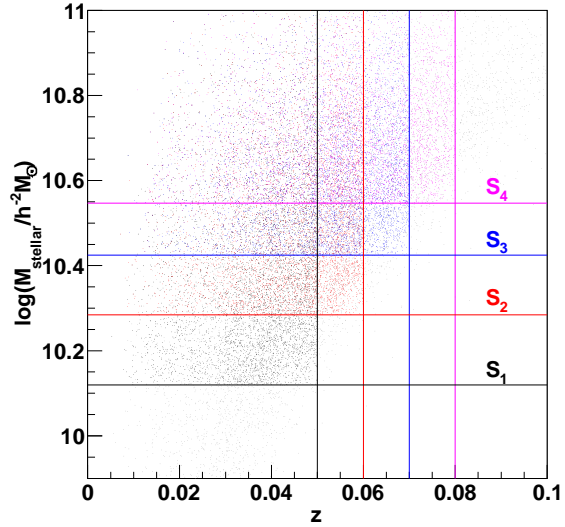
The 6dFGS  $J$ -selected sample used in this paper contains (after completeness cuts) 76 833 galaxies selected with  $9.8 \leq J \leq 13.75$ . We chose the  $J$ -band because it has the highest signal-to-noise of the three 2MASS bands. While there is slightly less extinction in the  $K$ -band compared to the  $J$ -band, for practical purposes, the  $J$ -band has better S/N because the night sky background glow is much less in the  $J$ - than in the  $K$ -band. The near infrared selection makes 6dFGS very reliable for stellar mass estimates.

The mean completeness of 6dFGS is 92 percent and the median redshift is  $z = 0.05$ . Completeness corrections are derived by normalising completeness-apparent magnitude functions so that, when integrated over all magnitudes, they equal the measured total completeness on a particular patch of sky. This procedure is outlined in the luminosity function evaluation of Jones et al. (2006) and also in Jones et al., (in prep). The original survey papers (Jones et al. 2004, 2005, 2009) describe in full detail the implementation of the survey and its associated online database.

The clustering in a galaxy survey is estimated relative to a random (unclustered) distribution which follows the same angular and redshift selection function as the galaxy sample itself. We base our random mock catalogue generation on the 6dFGS luminosity function (Jones et al. 2006), where we use random numbers to pick volume-weighted redshifts and luminosity function-weighted absolute magnitudes. We then test whether the redshift-magnitude combination falls within the 6dFGS  $J$ -band faint and bright apparent magnitude limits ( $9.8 \leq J \leq 13.75$ ). We assigned a  $b_J - r_F$  colour to each random galaxy using the redshift-  $b_J - r_F$  colour relation measured in the data and used these to derive stellar masses for the random galaxies using the same technique as for the actual galaxies (see section 2.1).

### 2.1 Stellar mass estimate and volume-limited sub-samples

To calculate the stellar mass for our dataset we use the stellar population synthesis results from Bruzual & Charlot



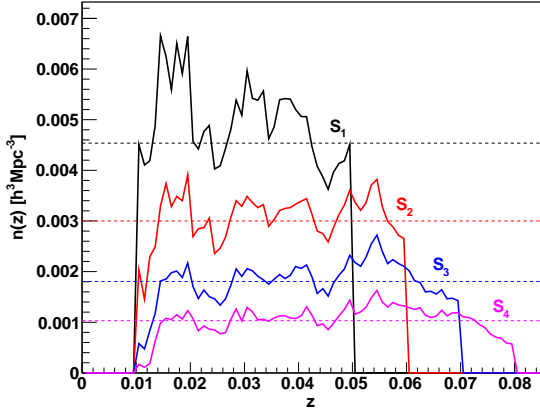
**Figure 1.** The distribution of 6dFGS galaxies in log stellar mass and redshift. The redshift and stellar mass cuts imposed to create the four volume-limited samples ( $S_1$  to  $S_4$  see Table 1) are shown by the coloured lines. All galaxies in the upper left quadrant created by the two correspondingly coloured lines are included in the volume-limited sub-samples. The plot shows a randomly chosen set of 20% of all galaxies.

(1993) together with a scaled Salpeter initial mass function (IMF) as reported in Bell & De Jong (2001)

$$\begin{aligned} \log_{10}(M_{\text{stellar}}/L_J) &= -0.57C_{b_J-r_F} + 0.48 \\ \log_{10}(L_J) &= (M_J^{\text{sun}} - M_J)/2.5 \\ \log_{10}(M_{\text{stellar}}) &= \log_{10}(M_{\text{stellar}}/L_J) + \log_{10}(L_J), \end{aligned} \quad (1)$$

with the 2MASS  $b_J - r_F$  colours,  $C_{b_J-r_F}$  the  $J$ -band absolute magnitude,  $M_J$  and the  $J$ -band absolute magnitude of the sun,  $M_J^{\text{sun}} = 3.70$  (Worthey 1994). The biggest uncertainty in stellar mass estimates of this type is the choice of the IMF. Assuming no trend in IMF with galaxy type, the range of IMFs presented in the literature cause uncertainties in the absolute normalisation of the stellar  $M/L$ -ratio of a factor of 2 in the near-infrared (Bell & De Jong 2001). The 6dFGS stellar mass function as well as a comparison of different stellar mass estimates is currently in preparation (Jones et al. in prep).

We create four volume-limited sub-samples in redshift and stellar mass. This is done by choosing an upper limit in redshift ( $z_{\text{max}}$ ) and then maximising the number of galaxies by choosing a lower limit in stellar mass ( $M_{\text{stellar}}^{\text{min}}$ ), meaning that every galaxy above that stellar mass will be detected in 6dFGS, if its redshift is below the redshift limit (see Figure 1). Because of the distribution in  $b_J - r_F$  colour, a clear cut in absolute magnitude does not correspond to a clear cut in stellar mass and therefore our samples are not perfectly volume-limited. We use the absolute magnitude limit which corresponds to a chosen redshift limit and derive a stellar mass limit, using the  $b_J - r_F$  colour corresponding to the 50% height of the  $b_J - r_F$  distribution ( $b_J - r_F(50\%) = 1.28$ ). We create four sub-samples ( $S_1$ - $S_4$ ) with upper redshift cuts at  $z_{\text{max}} = 0.05, 0.06, 0.07$  and  $0.08$  and with the mean log stel-



**Figure 2.** Galaxy density as a function of redshift for the four different volume-limited sub-samples  $S_1$ - $S_4$ . All samples follow an approximately constant number density indicated by the dashed lines and listed in the last column of Table 1.

lar masses ranging from  $\log_{10}(M_{\text{stellar}}/h^{-2} M_{\odot}) = 10.41$  to 10.73. We also include a low redshift cut-off at  $z_{\text{min}} = 0.01$ . Since the stellar mass distributions for these samples overlap, especially for the higher stellar mass samples, the results are correlated to some extent (see Figure 1). All our sub-samples are summarised in Table 1. Figure 2 shows the galaxy density as a function of redshift for the four sub-samples. The roughly constant number density with redshift shows that our sub-samples are close to volume-limited.

### 3 DATA ANALYSIS

We measure the separation between galaxies in our survey along the line of sight ( $\pi$ ) and perpendicular to the line of sight ( $r_p$ ) and count the number of galaxy pairs on this two-dimensional grid. We do this for the 6dFGS data catalogue, a random catalogue with the same selection function, and a combination of data-random pairs. We call the pair-separation distributions obtained from this analysis step  $DD(r_p, \pi)$ ,  $RR(r_p, \pi)$  and  $DR(r_p, \pi)$ , respectively. In the analysis we used 30 random catalogues with the same size as the real data catalogue and average  $DR(r_p, \pi)$  and  $RR(r_p, \pi)$ . The random mocks are sampled from the 6dFGS luminosity function (Jones et al. (2006) and also Jones et al, in prep.), and hence they contain the same evolution of luminosity with redshift that we see in 6dFGS itself. The redshift-space correlation function is then given by the Landy & Szalay (1993) estimator:

$$\xi(r_p, \pi) = 1 + \frac{DD(r_p, \pi)}{RR(r_p, \pi)} \left( \frac{n_r}{n_d} \right)^2 - 2 \frac{DR(r_p, \pi)}{RR(r_p, \pi)} \left( \frac{n_r}{n_d} \right), \quad (2)$$

where the ratio  $n_r/n_d$  is given by

$$\frac{n_r}{n_d} = \frac{\sum_i^{N_r} w_i}{\sum_j^{N_d} w_j} \quad (3)$$

and the sums go over all random ( $N_r$ ) and data ( $N_d$ ) galaxies. Here we employ a completeness weighting,  $w_i$ , where we weight each galaxy by the inverse sky- and magnitude completeness at its area of the sky (Jones et al., in prep.).

From the two-dimensional correlation function,  $\xi(r_p, \pi)$ , we calculate the projected correlation function

$$w_p(r_p) = 2 \int_0^{\pi^{\text{max}}} d\pi \xi(r_p, \pi), \quad (4)$$

where we bin  $\xi(r_p, \pi)$  in 30 logarithmic bins from 0.1 to  $100h^{-1}$  Mpc in  $r_p$  and  $\pi$ . The upper integration limit in eq. 4 was chosen to be  $\pi^{\text{max}} = 50h^{-1}$  Mpc for all sub-samples. We tested different values for  $\pi^{\text{max}}$ , changing it between 20 and  $90h^{-1}$  Mpc without any significant effect to  $r_0$  or  $\gamma$  if the errors on  $w_p(r_p)$  are adjusted accordingly. The error on scales larger than  $r_p = 20$  is very large and hence the contribution of these scales to the fit is small.

To derive a covariance matrix for the projected correlation function we use the method of jack-knife re-sampling our galaxy samples. First we divide the dataset into  $N = 400$  subsets, selected in R.A. and Dec. Each re-sampling step excludes one subset before calculating the correlation function. The covariance matrix is then given by

$$C_{ij} = \frac{(N-1)}{N} \sum_{k=1}^N \left[ w^k(r_p^i) - \bar{w}(r_p^i) \right] \left[ w^k(r_p^j) - \bar{w}(r_p^j) \right], \quad (5)$$

where  $w^k(r_p^i)$  is the projected correlation function estimate at separation  $r_p^i$  with the exclusion of subset  $k$  and  $\bar{w}(r_p^i)$  is the mean.

We also note that wide-angle effects can be neglected in this analysis, since we are interested in small-scale clustering (see Beutler et al. (2011) and Beutler et al. (2012) for a detailed investigation of wide-angle effects in 6dFGS).

### 3.1 Fibre proximity limitations

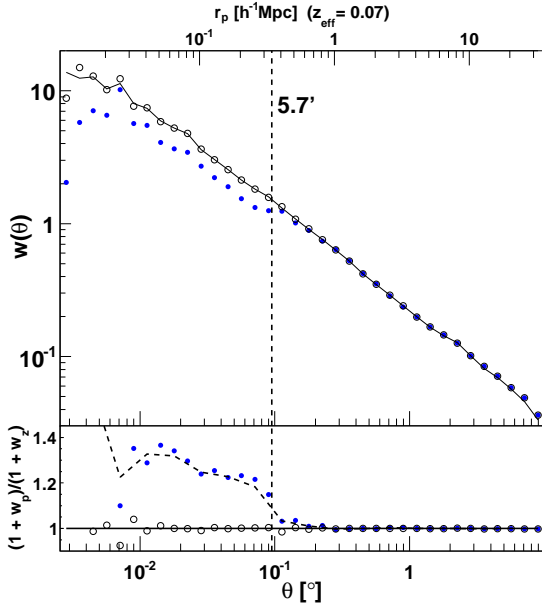
The design of the 6dF instrument does not allow fibres to be placed closer than 5.7 arcmin (Jones et al. 2004), which corresponds to a distance of  $r_p \approx 0.3h^{-1}$  Mpc at redshift  $z = 0.07$ . This limitation is relaxed in 6dFGS where about 70% of the survey area has been observed multiple times. However, for the remaining 30% we have to expect to miss galaxy pairs with a separation smaller than 5.7 arcmin.

Figure 3 shows the angular correlation function for the 6dFGS redshift catalogue ( $w_z$ , blue data points) and the target catalogue ( $w_p$ , solid black line). While the target catalogue contains 104 785 galaxies, the redshift catalogue of galaxies which have a  $J$ -band,  $b_J$ -band and  $r_F$ -band magnitude contains 76 833. The dashed line indicates the angular scale of the fibres. The two angular correlation functions agree on scales  $\theta > 0.1^\circ$ , with the redshift catalogue falling below the target catalogue at lower scales.

In the lower panel of Figure 3 we show the ratio  $(1 + w_p)/(1 + w_z)$ . The dashed line shows a spline fit to the blue data points, which then can be used to up-weight galaxy pairs with small angular separations and correct for the fibre proximity effect (Hawkins et al. 2003). This weighting is additional to the completeness weighting  $w_i$  we introduced earlier. While the completeness weighting is applied to single data- as well as random galaxies, the fibre proximity weighting is only applied to data galaxy pairs. Applying this weight to the 6dFGS redshift catalogue results in the open data points in Figure 3 which now agree very well with the target catalogue.

**Table 1.** Summary of the different volume-limited sub-samples,  $S_1$ - $S_4$ , used in this analysis. The effective redshift,  $z_{\text{eff}}$  and effective stellar mass are calculated as the mean of all galaxy pairs contributing to correlation function bins between  $0.1$  and  $40h^{-1}$  Mpc and the error is the standard deviation from the mean. The galaxy density  $n_g$  is calculated as the number of galaxies  $N$  divided by the co-moving sample volume, which is calculated from the maximum redshift  $z_{\text{max}}$ .

sample	$z_{\text{max}}$	$\log_{10} \left( \frac{M_{\text{stellar}}^{\text{min}}}{h^{-2} M_{\odot}} \right)$	$\langle z \rangle$	$\log_{10} \left( \frac{\langle M_{\text{stellar}} \rangle}{h^{-2} M_{\odot}} \right)$	$N$	$n_g [h^3 \text{Mpc}^{-3}]$
$S_1$	0.05	10.12	$0.0369 \pm 0.0010$	$10.410 \pm 0.010$	24 644	$4.536 \times 10^{-3}$
$S_2$	0.06	10.28	$0.0453 \pm 0.0011$	$10.532 \pm 0.010$	27 999	$3.001 \times 10^{-3}$
$S_3$	0.07	10.42	$0.0521 \pm 0.0012$	$10.635 \pm 0.018$	26 584	$1.806 \times 10^{-3}$
$S_4$	0.08	10.55	$0.0585 \pm 0.0020$	$10.730 \pm 0.013$	22 497	$1.030 \times 10^{-3}$



**Figure 3.** The angular correlation function of 6dFGS (blue data points) as well as the target catalogue (black solid line). The deviation between the two angular correlation functions is caused by fibre proximity limitations. The lower panel shows the ratio of the two correlation functions, which can be used as a weight to correct for the fibre proximity limitations (see section 3.1). The open data points show the corrected 6dFGS angular correlation function which is in very good agreement with the target catalogue.

#### 4 POWER LAW FITS

The projected correlation function can be related to the real-space correlation function  $\xi(r)$ , using (Davis & Peebles 1982)

$$\begin{aligned} w_p(r_p) &= 2 \int_0^\infty dy \xi \left[ (r_p^2 + y^2)^{1/2} \right] \\ &= 2 \int_{r_p}^\infty r dr \xi(r) (r^2 - r_p^2)^{-1/2}. \end{aligned} \quad (6)$$

If the correlation function is assumed to follow a power law,  $\xi(r) = (r/r_0)^\gamma$ , with the clustering amplitude  $r_0$  and the power law index  $\gamma$ , this can be written as

$$w_p(r_p) = r_p \left( \frac{r_p}{r_0} \right)^{-\gamma} \Gamma \left( \frac{1}{2} \right) \Gamma \left( \frac{\gamma-1}{2} \right) / \Gamma \left( \frac{\gamma}{2} \right), \quad (7)$$

**Table 2.** Best fitting parameters,  $r_0$  and  $\gamma$  for power law fits to the projected correlation functions  $w_p(r_p)$  of the four volume-limited 6dFGS sub-samples ( $S_1$ - $S_4$ ). The fitting range in all cases is  $0.1 < r_p < 40 h^{-1}$  Mpc with 24 bins and 2 free parameters. The last column shows the reduced  $\chi^2$  indicating the goodness of the fit.

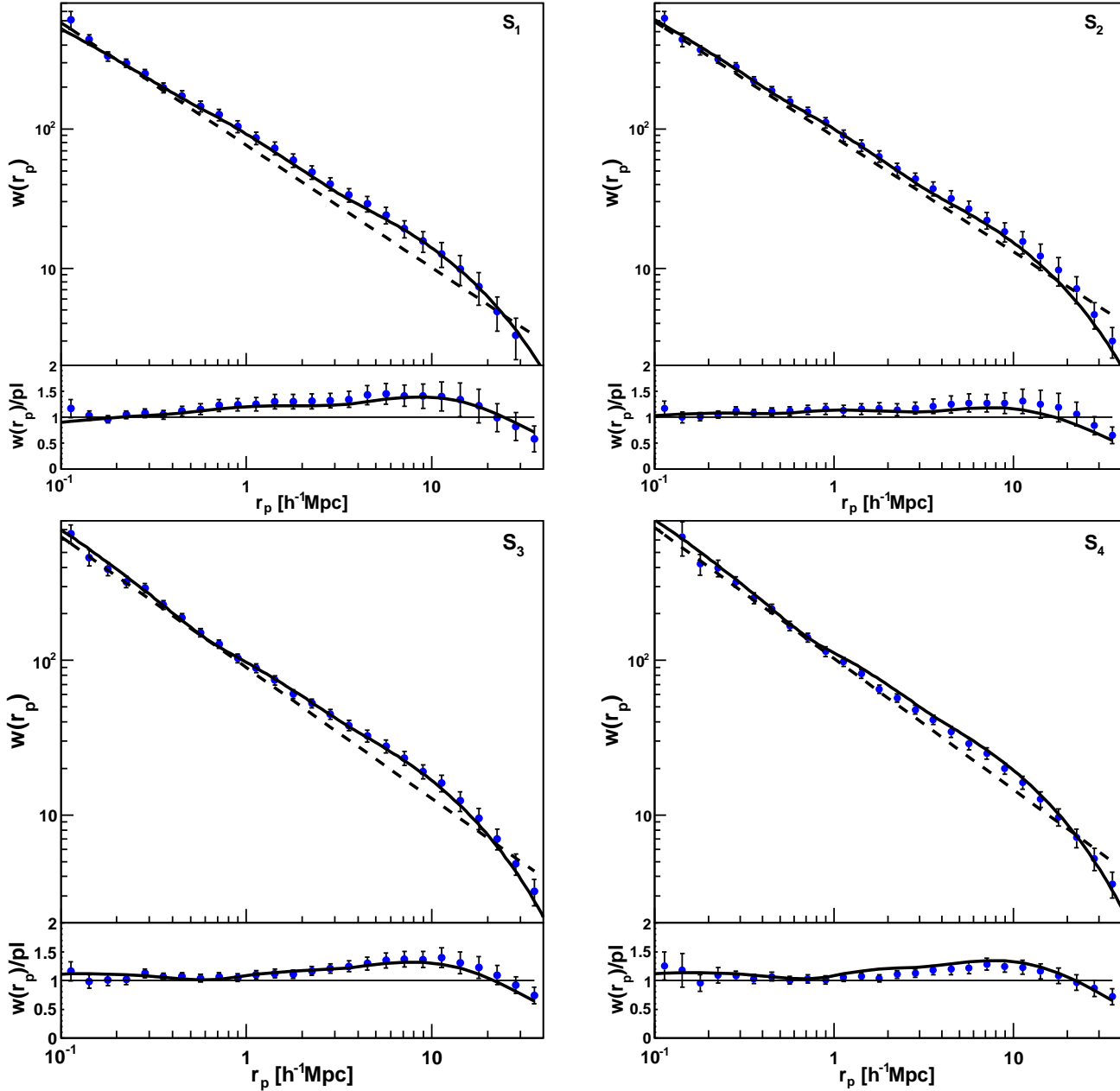
sample	$r_0 [h^{-1} \text{Mpc}]$	$\gamma$	$\chi^2/\text{d.o.f.}$
$S_1$	$5.14 \pm 0.23$	$1.849 \pm 0.025$	$23.5/(24-2) = 1.07$
$S_2$	$5.76 \pm 0.17$	$1.826 \pm 0.019$	$19.5/(24-2) = 0.89$
$S_3$	$5.76 \pm 0.16$	$1.847 \pm 0.019$	$32.6/(24-2) = 1.48$
$S_4$	$6.21 \pm 0.17$	$1.846 \pm 0.019$	$35.8/(24-2) = 1.63$

with  $\Gamma$  being the Gamma-function. Using this equation we can infer the best-fit power law for  $\xi(r)$  from  $w_p(r_p)$ .

Table 2 summarises the results of the power law fits to the four 6dFGS sub-samples. The best fitting power laws are also included in Figure 4 together with the four projected correlation functions. We set the fitting range to be  $0.1 < r_p < 40 h^{-1}$  Mpc which includes 24 bins. Although 6dFGS has very good statistics at scales smaller than  $0.1 h^{-1}$  Mpc we do not use them for our fits, since at such scales the fibre proximity correction becomes more than 30% (see section 3.1 and Figure 3).

The value of the clustering amplitude,  $r_0$ , increases from  $5.14$  to  $6.21 h^{-1}$  Mpc with increasing stellar mass, while there doesn't seem to be a clear trend in  $\gamma$ , which varies around the value of  $\gamma = 1.84$ . The reduced  $\chi^2$  in the last column of Table 2 indicates a good fit to the data for the first two sub-samples but grows to  $\chi^2/\text{d.o.f.} = 1.48$  and  $1.63$  for the high stellar mass sub-samples, indicating deviations from a power law. The lower panels of Figure 4 show the different projected correlation functions divided by the best fitting power law (blue data points). Here we can see that the deviations show systematic patterns. Such effects could be related to the strong correlations between bins in the correlation function. Nevertheless, these patterns can be addressed with a full HOD analysis, which we will pursue in the next section. We will compare the result of our power law fits with other studies in a more detailed discussion in section 8.

We remark at this point, that there is no satisfying theoretical model that predicts a power law behaviour of the correlation function, and hence the motivation of such a fit can only be empirical. While alternative approaches (e.g. the HOD model) rely on assumptions about the clustering behaviour of galaxies and dark matter, they are physically



**Figure 4.** The projected correlation function for the four different volume-limited 6dFGS sub-samples ( $S_1$ - $S_4$ ). The dashed black lines show the best fitting power laws (see Table 2) while the solid black lines show the best fitting HOD models, derived by fitting the data between  $0.1 \leq r_p \leq 40 h^{-1}$  Mpc. The lower panel shows the data and HOD models divided by the best fitting power laws.

motivated and hence allow us to learn more about galaxy clustering than a pure empirical power law fit.

## 5 THEORY: HALO OCCUPATION DISTRIBUTION

The Halo Occupation Distribution (HOD) model describes the relation between galaxies and mass in terms of the probability distribution  $p(N|M)$  that a halo of virial mass  $M$  contains  $N$  galaxies of a given type. Knowing how galaxies

populate dark matter halos, we can use a dark matter correlation function and infer the galaxy correlation function. We use CAMB (Lewis & Bridle 2002) to derive a model matter power spectrum which we turn into a correlation function using a Hankel transform

$$\xi(r) = \frac{1}{2\pi^2} \int_0^\infty dk P(k) k^2 \frac{\sin(kr)}{kr}. \quad (8)$$

The underlying cosmological model is fixed to  $(\Omega_b h^2, \Omega_c h^2, n_s, \sigma_8) = (0.02227, 0.1116, 0.966, 0.8)$  as reported in Komatsu et al. (2011) (we set  $\sigma_8 = 0.9$  for one special

case). We have to be aware of the fact that the fitted HOD parameters depend somewhat on the assumed values of  $\Omega_m$  and  $\sigma_8$  and hence the absolute values of the HOD parameters could be biased, if the assumed cosmology is wrong. However in this study we focus on the relative HOD parameters for different stellar mass selected sub-samples, which is fairly robust against such uncertainties.

Here we employ an analytic HOD methodology that is similar to that of Zheng, Coil & Zehavi (2007); Blake, Collister & Lahav (2008) and Zehavi et al. (2010). We utilise analytic approximations for the halo mass function (Tinker et al. 2008), the biased clustering of halos (Tinker et al. 2005) and the nonlinear dark matter power spectrum (Smith et al. 2003). The profile of dark matter within halos is well described by the NFW profile (Navarro, Frenk & White 1996) parameterized by the the concentration-mass relation (Duffy et al. 2008)

$$c(M, z) = 6.71 \left( \frac{M}{M_{\text{pivot}}} \right)^{-0.091} (1+z)^{-0.44} \quad (9)$$

with  $M_{\text{pivot}} = 2 \times 10^{12} h^{-1} M_\odot$ . We tried replacing eq. 9 with the form suggested by Bullock et al. (1999), and found that the best-fitting HOD parameters changed by much less than the statistical errors.

### 5.1 HOD framework and formalism

In the HOD parametrisation it is common to separate the clustering contributions from the most massive galaxies, which are assumed to sit in the halo centre, from satellite galaxies. This picture of how galaxies populate halos is supported by hydrodynamic simulations (e.g. Berlind et al. 2003; Simha et al. 2009) and semi-analytic models (e.g. White & Frenk 1991; Kauffmann, White & Guiderdoni 1993; Croton et al. 2006; Bower et al. 2006). The mean central and satellite number density of galaxies that populate dark matter halos of mass  $M$  is (Zheng et al. 2005)

$$\langle N_c(M) \rangle = \begin{cases} 0 & \text{if } M < M_{\text{min}} \\ 1 & \text{if } M \geq M_{\text{min}} \end{cases}, \quad (10)$$

$$\langle N_s(M) \rangle = \left( \frac{M}{M_1} \right)^\alpha,$$

where  $M_{\text{min}}$  is the minimum dark matter halo mass which can host a central galaxy,  $M_1$  corresponds to the mass of halos that contain, on average, one additional satellite galaxy ( $\langle N_s(M_1) \rangle = 1$ ) and  $\alpha$  sets the rate at which halos accumulate satellites when growing in mass. In very massive halos the number of satellites is proportional to halo mass  $M$  to the power of  $\alpha$ . The total HOD number is given by

$$\langle N_t(M) \rangle = \langle N_c(M) \rangle [1 + \langle N_s(M) \rangle], \quad (11)$$

so that a dark matter halo can only host a satellite galaxy if it contains already a central galaxy. In our model we assume a step like transition from  $\langle N_c(M) \rangle = 0$  to  $\langle N_c(M) \rangle = 1$ . In reality this is more likely to be a gradual transition with a certain width  $\sigma_{\log M}$  (More et al. 2009). To account for this, other studies (e.g. Zehavi et al. 2010) modify the HOD parametrisation of the central halo term to

$$\langle N_c(M) \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log_{10}(M) - \log_{10}(M_{\text{min}})}{\sigma_{\log M}} \right) \right], \quad (12)$$

which turns into eq. 10 for the case  $\sigma_{\log M} = 0$ . For our data we found that the reduction in  $\chi^2$  obtained from fits including  $\sigma_{\log M}$  as an additional parameter is not big enough to justify this parameterisation. For our largest sub-sample (S2) we found  $\Delta\chi^2 = -0.78$ , while a new parameter would be justified when  $\Delta\chi^2 < -2$  following the Akaike information criterion (Akaike et al. 1974).

There are also higher parameter models for the satellite fraction such as

$$\langle N_s(M) \rangle = \left( \frac{M - M_0}{M_1} \right)^\alpha, \quad (13)$$

where  $M_0$  is the minimum halo mass at which satellites can exist. Again we tested this model and found that the reduction in  $\chi^2$  does not justify this additional parameter. Hence we chose the three parameter model ( $M_1, \alpha, M_{\text{min}}$ ) of eq 10<sup>1</sup>.

Within the halo model we can account separately for the clustering amplitude of galaxies which sit in the same dark matter halo (one halo term) and galaxies which sit in different dark matter halos (two halo term). At small scales the clustering will be dominated by the one halo term and at large scales it will be dominated by the two halo term. For the correlation function this can be written as

$$\xi(r) = [1 + \xi_{1h}(r)] + \xi_{2h}(r). \quad (14)$$

where  $\xi_{1h}(r)$  and  $\xi_{2h}(r)$  represent the one halo and two halo terms respectively.

### 5.2 The 1-halo term, $\xi_{1h}(r)$

We separate the 1-halo term into contributions from central-satellite galaxy pairs and satellite-satellite galaxy pairs. The central-satellite contribution is given by

$$\xi_{1h}^{c-s}(r) = \frac{2}{n_g^2} \int_{M_{\text{vir}}(r)}^\infty dM \frac{dn(M)}{dM} N_c(M) N_s(M) \frac{\rho(r, M)}{M}, \quad (15)$$

where  $n_g$  is the galaxy number density and  $\rho(r, M)$  is the halo density profile. The lower limit for the integral is the virial mass  $M_{\text{vir}}(r)$  corresponding to the virial separation  $r_{\text{vir}}$ .

The satellite-satellite contribution is usually given by a convolution of the halo density profile with the halo mass function. Here we calculate this term in k-space since a convolution then turns into a simple multiplication

$$P_{1h}^{s-s}(k) = \frac{1}{n_g^2} \int_0^\infty dM \frac{dn(M)}{dM} N_c(M) N_s^2(M) \lambda(k, M)^2, \quad (16)$$

where  $\lambda(k, M)$  is the normalised Fourier transform of the halo density profile  $\rho(r, M)$ . The real-space expression of eq. (16) is

$$\xi_{1h}^{s-s}(r) = \frac{1}{2\pi^2} \int_0^\infty dk P_{1h}^{s-s}(k) k^2 \frac{\sin(kr)}{kr}, \quad (17)$$

which can then be combined with the central-satellite contribution to obtain the 1-halo term

$$\xi_{1h}(r) = \xi_{1h}^{c-s}(r) + \xi_{1h}^{s-s}(r) - 1. \quad (18)$$

<sup>1</sup>  $M_{\text{min}}$  is fixed by the number density

### 5.3 The 2-halo term, $\xi_{2h}(r)$

The 2-halo term,  $\xi_{2h}(r)$ , can be calculated from the dark matter correlation function since on sufficiently large scales the galaxy and matter correlation function are related by a constant bias parameter. We calculate the two halo term in Fourier space as

$$P_{2h}(k, r) = P_m(k) \times \left[ \int_0^{M_{\text{lim}}(r)} dM \frac{dn(M)}{dM} b_h(M, r) \frac{N_t(M)}{n_g(r)} \lambda(k, M) \right]^2, \quad (19)$$

where  $P_m(k)$  is the non-linear model power spectrum from CAMB including halo fit (Smith et al. 2003) and  $M_{\text{lim}}(r)$  is the halo mass limit for which we can find galaxy pairs with a separation larger than  $r_{\text{vir}}$ . To fix the value of  $M_{\text{lim}}(r)$  we use the number density matching model in the appendix of Tinker et al. (2005).  $b_h(M, r)$  is the scale dependent halo bias at separation  $r$  for which we assume the following model (Tinker et al. 2005)

$$b_h^2(M, r) = b^2(M) \frac{[1 + 1.17\xi_m(r)]^{1.49}}{[1 + 0.69\xi_m(r)]^{2.09}}, \quad (20)$$

where  $\xi_m(r)$  is the non-linear matter correlation function and  $b(M)$  being the bias function of Tinker et al. (2005).

### 5.4 Derived quantities

Since the HOD model is directly based on a description of dark matter clustering and its relation to galaxy clustering, an HOD model can tell us much more about a galaxy population, than just the two parameters  $M_1$  and  $\alpha$ .

For  $\lambda(k, M) \rightarrow 1$  (which corresponds to large separations  $r$  in real-space) the two halo term simplifies to

$$P_{2h}(k, r) \approx b_{\text{eff}}^2 P_m(k), \quad (21)$$

where the effective bias is the galaxy number weighted halo bias factor (Tinker et al. 2005)

$$b_{\text{eff}} = \frac{1}{n_g} \int_0^\infty dM \frac{dn(M)}{dM} b_h(M) N_t(M). \quad (22)$$

We can also ask what is the average group dark matter halo mass for a specific set of galaxies (often called host-halo mass). Such a quantity can be obtained as

$$M_{\text{eff}} = \frac{1}{n_g} \int_0^\infty dM \frac{dn(M)}{dM} M N_t(M), \quad (23)$$

which represents a weighted sum over the halo mass function with the HOD number as a weight.

The averaged ratio of satellite galaxies to the total number of galaxies is given by

$$f_s = \frac{\int_0^\infty dM \frac{dn(M)}{dM} N_s(M)}{\int_0^\infty dM \frac{dn(M)}{dM} N_c(M) [1 + N_s(M)]} \quad (24)$$

and the central galaxy ratio is given by  $f_c = 1 - f_s$ .

## 6 HOD PARAMETER FITS

To compare the HOD models to the data we use the same fitting range as for the power law fits earlier ( $0.1 < r_p <$

$40 h^{-1} \text{ Mpc}$ ). The free parameters of the fit are  $M_1$  and  $\alpha$  as described in the last section and hence we have the same number of free parameters as for the power law fits.

All the fitting results are summarised in Table 3 and Figure 4. The reduced  $\chi^2$  in the last column of Table 3 indicates a good fit to the data in all cases and the ratio of data to best fitting power law in the lower panels of Figure 4 shows that the HOD model reproduces the double peak structure present in the data. We also note that the reduced  $\chi^2$  in case of the HOD fits is uniformly lower than for the power law fits indicating a better fit to the data for all sub-samples.

Figure 5 shows the 2D probability distributions in  $M_1$  and  $\alpha$  for the four different volume-limited sub-samples. There is a strong trend of increasing  $M_1$  with stellar mass and a weaker but still significant trend of increasing  $\alpha$  with increasing stellar mass. This indicates that dark matter halos that host a central galaxy with higher stellar mass have their first satellite on average at a larger dark matter halo mass. However the number of satellites increases more steeply with halo mass for samples with higher stellar mass. This trend indicates that halos of higher mass have greater relative efficiency at producing multiple satellites. Similar trends were found in sub-samples of SDSS galaxies by Zehavi et al. 2005b. As a test for the sensitivity of the fitting results to the upper fitting limit we performed a fit to sub-sample  $S_2$  with the fitting range  $0.1 < r_p < 20 h^{-1} \text{ Mpc}$ . All parameters agree within  $1\sigma$  with the fit to the larger fitting range. We called this fit  $S'_2$  and included it in Table 3.

Table 3 also includes derived parameters like the minimum dark matter halo mass,  $M_{\text{min}}$ , the satellite fraction,  $f_s$ , the effective dark matter halo mass,  $M_{\text{eff}}$  and the effective galaxy bias,  $b_{\text{eff}}$ . The errors on these parameters are calculated as the 68%-confidence level of their 1D probability distribution. The effective dark matter halo mass seems to be almost constant for all samples while the minimum dark matter halo mass  $M_{\text{min}}$  increases with stellar mass. The satellite fraction  $f_s$  decreases with increasing stellar mass, indicating that galaxies with high stellar mass have a higher probability to be central galaxies. The increasing effective galaxy bias indicates that galaxies with higher stellar mass are more strongly clustered and hence reside in high density regions of the Universe. The increasing effective galaxy bias is well described by the following form:

$$b(M_{\text{stellar}}) = (1.05 + M_{\text{stellar}}/M_*) \times (0.8/\sigma_8) \quad (25)$$

with  $M_* = 1.18 \times 10^{11} h^{-2} M_\odot$ . We compare this function to the measurements in Figure 6. The absolute stellar masses are subject to significant uncertainties and different methods to derive stellar masses can come to very different conclusions. However, most stellar mass estimates are related by a simple constant offset, and relation 25 can be scaled accordingly. The equation above is only valid for the stellar mass range probed in this analysis ( $M_{\text{stellar}} = 2.6 - 5.4 \times 10^{10} h^{-2} M_\odot$ ) since the underlying dynamics are most likely not captured in eq. 25.

We will discuss the implications of all these results in the next sections. First we will derive 6dFGS mock samples using different semi-analytic models. We will then compare the predictions from these semi-analytic models to our data followed by a comparison to other studies.

**Table 3.** Summary of the best fitting parameters for HOD fits to the four 6dFGS sub-samples ( $S_1$ - $S_4$ ) with the fitting range  $0.1 < r_p < 40h^{-1}$  Mpc. The last column shows the reduced  $\chi^2$  derived from a fit to 24 bins with 2 free parameters and indicates good fits for all sub-samples. Error bars on the HOD parameters correspond to  $1\sigma$ , derived from the marginalised distributions. The sample  $S'_2$  uses a different fitting range of  $0.1 < r_p < 20h^{-1}$  Mpc for the largest of our sub-samples,  $S_2$ . Furthermore we include a special fit to  $S_2$  labeled  $S_2^{\sigma_8=0.9}$ , where we change our standard assumption of  $\sigma_8 = 0.8$  to  $\sigma_8 = 0.9$ . The satellite fraction  $f_s$ , the effective dark matter halo mass  $M_{\text{eff}}$  and the effective galaxy bias  $b_{\text{eff}}$  are derived parameters (see section 5.4). The bias depends on our initial assumption of  $\sigma_8$  and hence this parameter should be treated as  $b_{\text{eff}} \times (\sigma_8/0.8)$ .

sample	$\log_{10} \left( \frac{M_1}{h^{-1}M_{\odot}} \right)$	$\alpha$	$\log_{10} \left( \frac{M_{\text{min}}}{h^{-1}M_{\odot}} \right)$	$f_s$	$\log_{10} \left( \frac{M_{\text{eff}}}{h^{-1}M_{\odot}} \right)$	$b_{\text{eff}}$	$\chi^2/\text{d.o.f.}$
$S_1$	$13.396 \pm 0.017$	$1.214 \pm 0.031$	$12.0478 \pm 0.0049$	$0.2106 \pm 0.0078$	$13.501 \pm 0.015$	$1.2704 \pm 0.0087$	$13.5/(24-2) = 0.61$
$S_2$	$13.568 \pm 0.013$	$1.270 \pm 0.027$	$12.2293 \pm 0.0043$	$0.1879 \pm 0.0062$	$13.532 \pm 0.012$	$1.3443 \pm 0.0063$	$16.6/(24-2) = 0.75$
$S_3$	$13.788 \pm 0.012$	$1.280 \pm 0.029$	$12.4440 \pm 0.0039$	$0.1578 \pm 0.0061$	$13.546 \pm 0.012$	$1.4002 \pm 0.0062$	$21.4/(24-2) = 0.97$
$S_4$	$14.022 \pm 0.011$	$1.396 \pm 0.033$	$12.6753 \pm 0.0032$	$0.1243 \pm 0.0058$	$13.617 \pm 0.012$	$1.5170 \pm 0.0077$	$24.4/(24-2) = 1.11$
$S_2^{\sigma_8=0.9}$	$13.605 \pm 0.013$	$1.193 \pm 0.023$	$12.2382 \pm 0.0048$	$0.2034 \pm 0.0077$	$13.649 \pm 0.014$	$1.2151 \pm 0.0065$	$19.6/(24-2) = 0.89$
$S'_2$	$13.568 \pm 0.013$	$1.282 \pm 0.027$	$12.2291 \pm 0.0042$	$0.1890 \pm 0.0074$	$13.542 \pm 0.010$	$1.3492 \pm 0.0065$	$19.1/(21-2) = 1.01$

## 7 SEMI-ANALYTIC MOCK CATALOGUES

To compare our result with theory, we derive 6dFGS mock catalogues from two different semi-analytical models (Croton et al. 2006; Bower et al. 2006), both based on the Millennium Simulation (Springel et al 2005) publicly available through the Millennium Simulation database<sup>2</sup>. Semi-analytic models are based on an underlying N-body simulation together with theoretically and observationally motivated descriptions of gas cooling, star formation and feedback processes.

The Millennium Simulation is a dark matter only N-body simulation which traces the hierarchical evolution of  $2160^3$  particles in a periodic box of  $500^3 h^{-3} \text{Mpc}^3$  from redshift  $z = 127$  to  $z = 0$ . The underlying cosmological model follows WMAP1 cosmology (Spergel et al 2003) given by a matter density of  $\Omega_m = \Omega_{dm} + \Omega_b = 0.25$ , a cosmological constant of  $\Omega_{\Lambda} = 0.75$ , a Hubble constant of  $H_0 = 75 \text{ km/s/Mpc}$ , a spectral index of  $n_s = 1$  and a r.m.s. of matter fluctuations in  $8h^{-1} \text{Mpc}$  spheres of  $\sigma_8 = 0.9$ . The individual particle mass of the simulation is  $8.6 \times 10^8 h^{-1} M_{\odot}$  and halos and sub-halos are identified from the spatial distribution of dark matter particles using a standard friends-of-friends algorithm and the SUBFIND algorithm (Springel et al 2001). All sub-halos are then linked together to construct the halo merger trees which represent the basic input of the semi-analytic models.

Here we are using the  $z = 0$  output of the Millennium Simulation. To ensure that the stellar masses are calculated in a consistent and comparable way, we implement the following procedure to derive the 6dFGS mock catalogues:

(i) We apply the 6dFGS  $K$ -band apparent magnitude limits of  $8.85 \leq K \leq 12.75$  to the full  $500h^{-3} \text{Mpc}^3$  simulation box. We have to use the  $K$ -band instead of the  $J$ -band, which is actually used in this analysis, because none of the semi-analytic models provide  $J$ -band magnitudes. However, the 6dFGS  $J$ -band and  $K$ -band samples have significant overlap, meaning that almost all galaxies which have a  $K$ -

band magnitude also have a  $J$ -band magnitude. We also account for sky- and magnitude incompleteness.

(ii) We re-calculate stellar masses for each galaxy using the technique described in section 2.1 but with the corresponding  $K$ -band relations (Bell & De Jong 2001), instead of  $J$ -band. The re-calculation of the stellar masses ensures that potential disagreement with our measurement is not caused by a different technique of deriving stellar masses or a different assumptions about the IMF.

(iii) We apply the same redshift and stellar mass limits to the semi-analytic catalogues, which we used to produce the four volume-limited samples in the 6dFGS dataset (see Table 1).

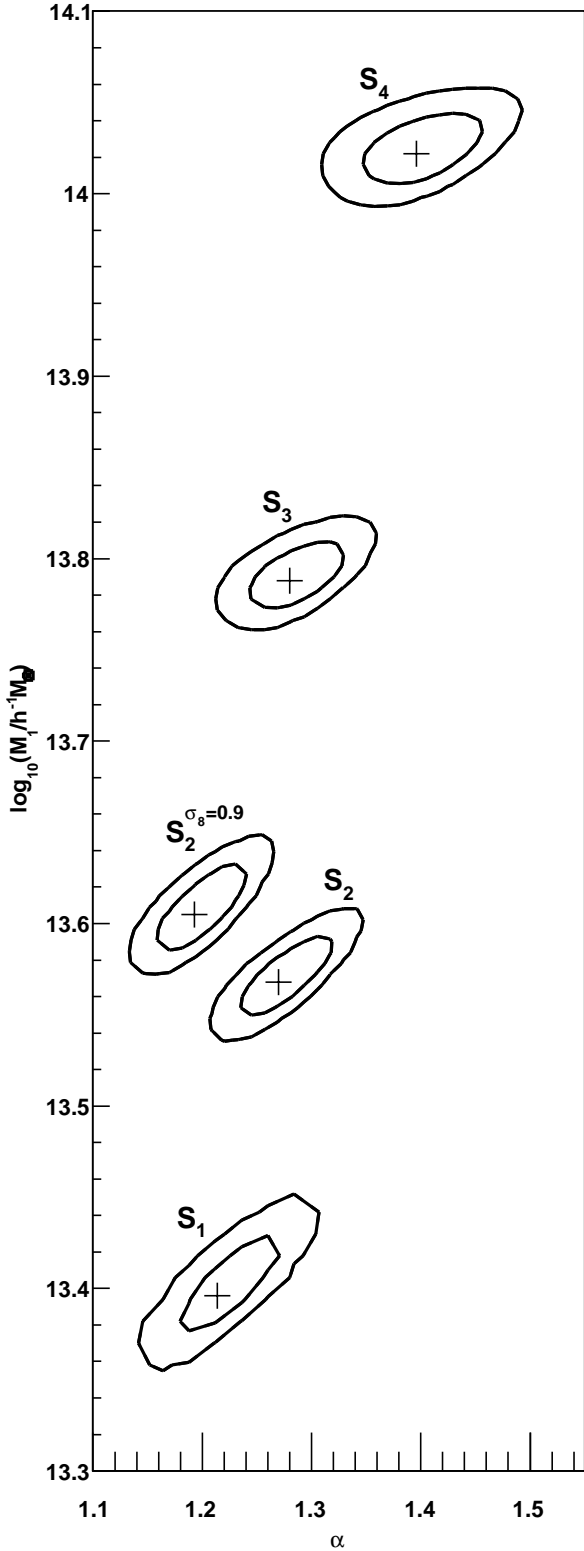
We found that all mock 6dFGS catalogues derived from these semi-analytic models contain fewer galaxies than the data sample (by about 40%). This could be related to the slightly different cosmology used in these simulations, which should have its largest impact on large clusters, which are sampled in 6dFGS. We are not attempting to correct for such differences in the cosmological model. The aim of this part of our analysis is to test the current predictive power of semi-analytic models.

Semi-analytic models are often grouped into “Durham models” and “Munich models”. The Bower et al. (2006) model belongs in the group of “Durham models” while the Croton et al. (2006) model belongs in the group of “Munich models”.

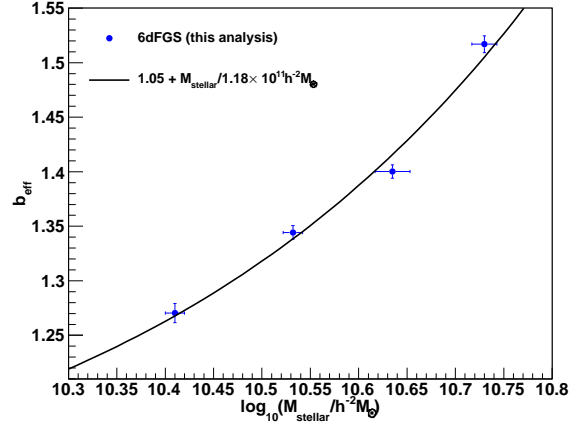
### 7.1 Durham models

In the Durham models, merger trees are produced following Helly et al. (2003) which are independent of those generated by Springel et al (2005). When the satellite galaxy falls below a certain distance to the central galaxy given by  $R_{\text{merge}} = r_c + r_s$ , where  $r_c$  and  $r_s$  are the half mass radii of the central and satellite galaxy, respectively, the satellite and central galaxy are treated as one. The largest of the galaxies contained within this new combined dark matter halo is assumed to be the central galaxy (Benson et al. 2002), whilst all other galaxies within the halo are satellites. The dynamical friction and tidal stripping which are present in such a system are modelled analytically. These models are

<sup>2</sup> <http://gavo.mpa-garching.mpg.de/MyMillennium/>  
<http://galaxy-catalogue.dur.ac.uk:8080/Millennium/>



**Figure 5.** Two dimensional probability distribution for  $\log_{10}(M_1)$  and  $\alpha$  for the HOD fits to the four 6dFGS volume-limited sub-sample ( $S_1$ - $S_4$ ). We include two fits for the largest sample  $S_2$ , where for  $S_2^{\sigma_8=0.9}$  we change our standard assumption of  $\sigma_8 = 0.8$  to  $\sigma_8 = 0.9$ . The parameters derived from the fits are summarised in Table 3. The best fitting values are marked with black crosses.



**Figure 6.** The effective galaxy bias (see eq. 22) as a function of  $\log$  stellar mass for 6dFGS (blue data points). The increase in galaxy bias with stellar mass can be described by eq. 25 which corresponds to the black line.

based on NFW density profiles for the central halo as well as the satellite halo, while galaxies are modelled as a disc plus spheroid.

## 7.2 Munich models

The Munich models are based on the original merger trees by Springel et al (2005). One of the key differences between these merger trees and the ones used in the Durham models is that the Munich models explicitly follow dark matter halos even after they are accreted onto larger systems, allowing the dynamics of satellite galaxies residing in the in-falling halos to be followed until the dark matter substructure is destroyed. The galaxy is then assigned to the most bound particle of the sub-halo at the last time the sub-halo could be identified.

In the Munich models, a two-mode formalism is adopted for active galactic nuclei (AGNs), wherein a high-energy, or “quasar” mode occurs subsequent to mergers, and a constant low-energy “radio” mode suppresses cooling flows due to the interaction between the gas and the central black hole (Croton et al. 2006). In the quasar model, accretion of gas onto the black hole peaks at  $z \sim 3$ , while the radio mode reaches a plateau at  $z \sim 2$ . AGN feedback is assumed to be efficient only in massive halos, with supernova feedback being more dominant in lower-mass halos.

The galaxies in the Croton et al. (2006) model can be of three different types: central galaxies (type 0), satellites of type 1 and satellites of type 2. Satellites of type 1 are associated with dark matter substructures, which usually refers to recently merged halos. Satellites of type 2 are instead galaxies whose dark matter halo has completely merged with a bigger halo and are not associated with a substructure. We treat both type 2 and type 1 as satellite galaxies.

## 7.3 Testing semi-analytic models

In their original paper Bower et al. (2006) compare the K-band luminosity function, galaxy stellar mass function,

and cosmic star formation rate with high-redshift observations. They find that their model matches the observed mass and luminosity functions reasonably well up to  $z \approx 1$ . Kitzbichler & White (2007) compare the magnitude counts in the  $b_J$ ,  $r_F$ ,  $I$  and  $K$ -bands, redshift distributions for  $K$ -band selected samples,  $b_J$ - and  $K$ -band luminosity functions, and galaxy stellar mass function from Croton et al. (2006) and the very similar model by De Lucia & Blaizot (2007) with high redshift measurements. They find that the agreement of these models with high-redshift observations is slightly worse than that found for the Durham models. In particular, they find that the Munich models tend to systematically overestimate the abundance of relatively massive galaxies at high redshift.

Snaith et al. (2011) compared four different semi-analytic models (De Lucia & Blaizot 2007; Bower et al. 2006; Bertone et al. 2007; Font et al. 2008), with observations and found that all models show a shallower, wider magnitude gap, between the brightest group galaxy and the second brightest, compared to observations.

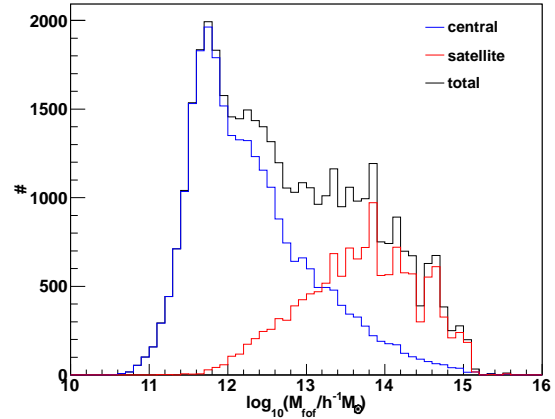
de la Torre et al. (2010) compared measurements of VVDS with the model by De Lucia & Blaizot (2007). They found that the model reproduced the galaxy clustering at  $z > 0.8$  as well as the magnitude counts in most bands. However the model failed in reproducing the clustering strength of red galaxies and the  $b_J - I$  colour distribution. The model tends to produce too many relatively bright red satellite galaxies, a fact that has been reported in other studies as well (over-quenching problem: Weinmann et al. 2006; Kimm et al. 2009; Liu et al. 2010).

## 8 DISCUSSION

### 8.1 Effective halo mass and satellite fraction

As is evident from Table 3, the 6dFGS galaxies sit in massive central dark matter halos and most of our galaxies are central galaxies in these halos with only a small fraction being satellite galaxies. Because of the way 6dFGS galaxies are selected, the large majority are red elliptical galaxies and hence our findings agree very well with previous studies, which also found that such galaxies are strongly clustered and therefore must reside in high density regions.

First we will discuss the effective dark matter halo mass, which represents the effective group or cluster mass for the distribution of galaxies (not the mass of the individual halos which host the galaxies). This quantity appears almost constant for our four different sub-samples. The increasing central galaxy mass ( $M_{\min}$ ) with stellar mass is offset by the decreasing satellite fraction, resulting in a fairly constant  $M_{\text{eff}}$ . While the HOD model allows us to derive only averaged parameters for each sample, semi-analytic models directly connect dark matter halo masses with single galaxies. Figure 7 shows the distribution of central and satellite galaxies as a function of friends-of-friends (fof) halo mass derived from the Bower et al. (2006) semi-analytic catalogue together with the 6dFGS selection criteria. The catalogue contains fewer satellites than central, but the satellites sit in very massive dark matter halos and hence have a significant impact on  $M_{\text{eff}}$ . This plot shows that most 6dFGS galaxies sit in  $10^{11} - 10^{12} h^{-1} M_{\odot}$  halos, while the satellites



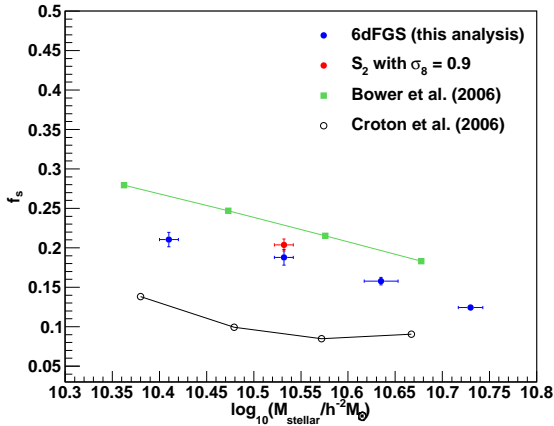
**Figure 7.** This plot shows the friends-of-friends halo mass  $\log_{10}(M_{\text{fof}})$  of the satellite and central galaxies from the Bower et al. (2006) semi-analytic catalogue with the 6dFGS selection. The bimodal distribution is caused by a dominant fraction of halos, in which the 6dFGS galaxies are the central galaxy and a smaller fraction with larger halo mass, in which the 6dFGS galaxies are satellite galaxies.

**Table 4.** Summary of parameters derived from different semi-analytic models. We impose the 6dFGS  $K$ -band apparent magnitude selection ( $J$ -band is not available for these semi-analytic models) as well as correct for incompleteness. The stellar masses are re-calculated using the technique described in section 2.1. To calculate the effective dark matter halo mass we used the friends-of-friends halo mass. The stellar mass and the effective halo mass are calculated as the mean of the distribution.

sample	$f_s$	$\log_{10} \left( \frac{\langle M_{\text{stellar}} \rangle}{h^{-2} M_{\odot}} \right)$	$\log_{10} \left( \frac{M_{\text{eff}}}{h^{-1} M_{\odot}} \right)$
(Croton et al. 2006)			
S1	0.138	10.38	13.400
S2	0.099	10.48	13.389
S3	0.085	10.57	13.507
S4	0.090	10.67	14.013
(Bower et al. 2006)			
S1	0.28	10.36	13.695
S2	0.25	10.47	13.661
S3	0.22	10.58	13.674
S4	0.18	10.68	13.777

sit in very massive groups and clusters of up to  $10^{15} h^{-1} M_{\odot}$ . While the median would be around  $10^{12} h^{-1} M_{\odot}$ , the effective mass  $M_{\text{eff}}$  from the HOD model is the mean of this distribution, which is pushed to very large values by the satellite fraction.

We calculated the effective halo mass  $M_{\text{eff}} = \langle M_{\text{fof}} \rangle$  for the four volume-limited sub-samples derived from the two semi-analytic models and summarise these results in Table 4. While the effective halo mass appears constant in case of the Bower et al. (2006) model, the Croton et al. (2006) model shows an increase with increasing stellar mass. These different behaviours for the two different semi-analytic mod-



**Figure 8.** Satellite fraction of the 6dFGS stellar mass sub-samples (blue) and the corresponding semi-analytic sub-samples based on the Millennium simulation (Croton et al. 2006; Bower et al. 2006). The red data point represents a fit to sample  $S_2$  which assumes  $\sigma_8 = 0.9$ , while the blue data points assume  $\sigma_8 = 0.8$ .

els are most likely connected to the different trend in the satellite fraction shown in Figure 8. While the Bower et al. (2006) model shows a constant decrease in the satellite fraction, the Croton et al. (2006) model reaches a constant at large stellar mass which causes the effective halo mass to rise. A lower satellite fraction at a fixed stellar mass means that at fixed halo mass the satellite galaxies are less massive (Weinmann et al. 2006; Kimm et al. 2009; Liu et al. 2010).

From Figure 8 we can see that both the Croton et al. (2006) and the Bower et al. (2006) model give slightly smaller stellar masses than measured in 6dFGS. This indicates that the semi-analytic catalogues contain too few bright galaxies at low redshift.

Figure 8 also includes the fit to sample  $S_2$  which assumes  $\sigma_8 = 0.9$  (red data point), which agrees with the assumptions of the Millennium Simulation, while the blue data points assume  $\sigma_8 = 0.8$ . A larger  $\sigma_8$  causes a larger satellite fraction which then causes a larger effective halo mass. This effect seems to bring our results closer to the Bower et al. (2006) model.

## 8.2 The $M_1 - M_{\min}$ scaling relation

The HOD analysis of early data from SDSS by Zehavi et al. (2005b) showed the relation,  $M_1 \approx 23 M_{\min}$ , between the mass of halos that on average host one additional satellite galaxy,  $M_1$  and the minimum dark matter halo mass to host a central galaxy,  $M_{\min}$ . This has been confirmed in subsequent studies ( $M_1 \approx 18 M_{\min}$  in Zheng, Coil & Zehavi 2007 and  $M_1 \approx 17 M_{\min}$  Zehavi et al. 2010). This relation implies that on average a halo hosting two galaxies of the type studied in their analysis has a mass  $\approx 23$  times the mass of a halo hosting only one galaxy of the same type.

Zehavi et al. (2010) also found that this scaling factor is somewhat smaller at the high luminosity end, corresponding to massive halos that host rich groups or clusters. This latter

trend likely reflects the relatively late formation of these massive halos, which leaves less time for satellites to merge onto central galaxies and thus lowers the satellite threshold  $M_1$ .

Theoretical studies of HODs in dark matter simulations (Kravtsov et al. 2004) and those predicted by Smoothed-particle hydrodynamics (SPH) and semi-analytic galaxy formation models (Zheng et al. 2005) reveal a similar relation with a scaling factor of  $\approx 20$ . The large gap between  $M_1$  and  $M_{\min}$  arises because in the low occupation regime, a more massive halo tends to host a more massive central galaxy, rather than multiple smaller galaxies (Berlind et al. 2003).

Abbas et al. (2010) did a similar study using data from the  $I$ -band selected VIMOS-VLT Deep Survey (VVDS) and found the ratio  $M_1/M_{\min}$  to be  $\approx 40 - 50$ . The galaxies in this study are at much higher redshift ( $z \approx 0.83$ ) compared to the SDSS galaxies. This result means that in order to begin hosting satellite galaxies, halos sampled by the VVDS survey need to accrete a larger amount of mass compared to SDSS halos.

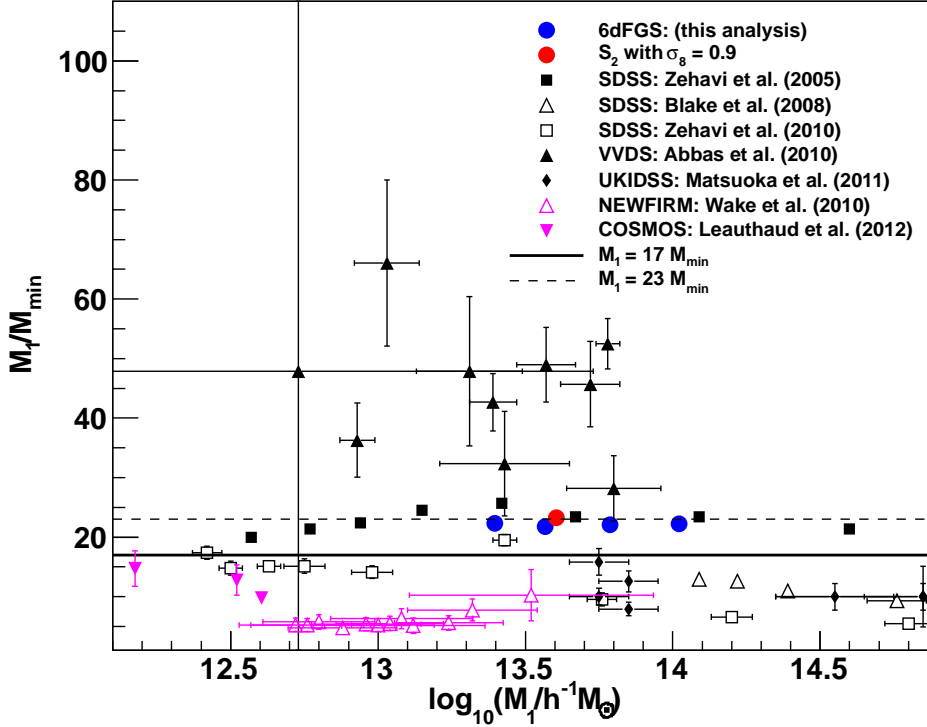
Wake et al. (2011) also studied high redshift galaxies ( $1.1 < z < 1.9$ ) in the NEWFIRM Medium Band Survey (NMBS) and found the ratio  $M_1/M_{\min}$  to be  $\approx 4 - 10$ . This, together with the result by Abbas et al. (2010) shows, that this relation is not as fundamental as originally thought, but strongly dependent on the type of halos probed in each analysis.

Leauthaud et al. (2012) studied galaxy clustering in the COSMOS survey using threshold stellar mass samples. Using the equations discussed in (Leauthaud et al. 2011) we can derive  $M_1$  and  $M_{\min}$  from their HOD parametrisation, which is included in Figure 9.

Matsuoka et al. (2011) analysed  $\sim 60\,000$  massive ( $\log_{10}(M_{\text{stellar}}/h^{-2}M_{\odot}) > 10.7$ ) galaxies from the UKIRT Infrared Deep Sky Survey (UKIDSS) and the SDSS II Supernova Survey. This analysis shows a very different clustering amplitude depending on whether the observed and theoretical number densities are matched up or not. This dependency makes a comparison of our results with their derived parameters very difficult. Nevertheless, the ratio  $M_1/M_{\min}$  in their analysis does not depend significantly on their initial assumptions and hence we included their results in Figure 9<sup>3</sup>. Their results indicate a lower  $M_1/M_{\min}$  ratio at large  $M_1$  consistent with Zehavi et al. (2010) and Blake, Collister & Lahav (2008).

In our analysis we found a scaling relation of  $M_1 \approx 22 M_{\min}$  (the exact values are  $22.29 \pm 0.39$ ,  $21.81 \pm 0.30$ ,  $22.08 \pm 0.28$  and  $22.22 \pm 0.25$  for  $S_1$ - $S_4$ , respectively). In Figure 9 we compare our results with other stellar mass selected samples (coloured data points) and with luminosity threshold selected samples (black data points). The ratio  $M_1/M_{\min}$  found in 6dFGS is in agreement with Zehavi et al. (2005b) and slightly larger than Zehavi et al. (2010). While Zehavi et al. (2005b) and Zehavi et al. (2010) study the same type of galaxies, Zehavi et al. (2005b) uses an HOD parametrisation very similar to ours, while Zehavi et al. (2010) uses a parameterisation based on 5 free parameters,

<sup>3</sup> We use their results, in which they match the observed and theoretical number densities, since this agrees with our method.



**Figure 9.** The relation between  $M_1$  and  $M_{\min}$  for the four different volume-limited 6dFGS samples (blue data points), compared to Zehavi et al. (2005b) (black solid squares), Blake, Collister & Lahav (2008) (black open triangles), Zehavi et al. (2010) (black open squares), Abbas et al. (2010) (black solid triangles), Matsuoka et al. (2011) (black open points), Wake et al. (2011) (magenta open triangles) and Leauthaud et al. (2012) (magenta solid triangles). Zehavi et al. (2005b) don't report errors on their parameters and the 6dFGS errors are smaller than the data points. All studies which investigate HOD as a function of stellar mass are coloured (including our analysis), while luminosity defined studies are in black.

suggesting that the differences might be related to the parameterisation.

### 8.3 Comparison to Mandelbaum et al. (2006)

An alternative method for probing the connection between stellar mass and halo mass is galaxy-galaxy weak lensing. Here we will compare our findings to Mandelbaum et al. (2006), who used weak lensing of  $\sim 350\,000$  galaxies from SDSS and looked at the dependence of the amplitude of the lensing signature as a function of galaxy type and stellar mass. They used stellar masses derived from the  $z$ -band magnitude and the ratio  $M/L_z$  from Kauffmann et al. (2003) with the assumption of a Kroupa (2001) IMF.

The HOD model employed in there analysis is given by

$$N_t(M) = N_c(M) + N_s(M) \quad (26)$$

with the central galaxy number given by eq. 10 while the satellite galaxies number is modelled by a step-like function

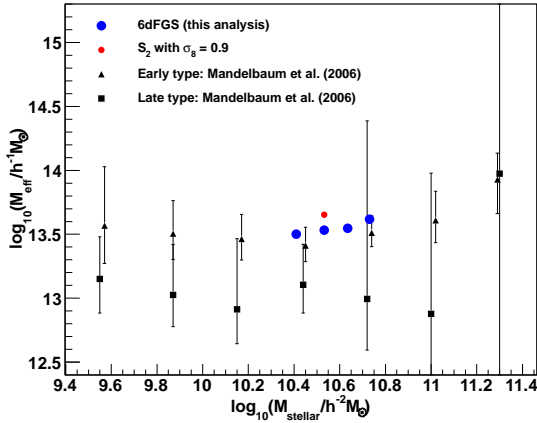
$$N_s(M) = \begin{cases} kM & \text{if } M \geq 3M_{\min} \\ \frac{kM^2}{3M_{\min}} & \text{if } M_{\min} \leq M < 3M_{\min} \\ 0 & \text{if } M < M_{\min}. \end{cases} \quad (27)$$

The normalisation constant  $k$  can be determined by matching the measured satellite fractions  $f_s^{\text{ma}}$  and is given

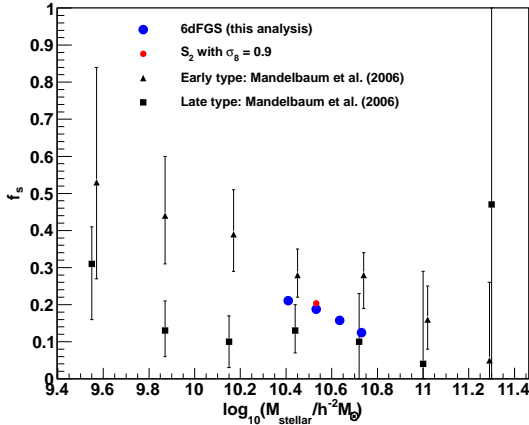
in appendix A. Here we have adjusted the nomenclature to the one used in our analysis. We also note that Mandelbaum et al. (2006) assumed  $\sigma_8 = 0.9$ , while we assumed  $\sigma_8 = 0.8$  for most of our fits.

From the equation above we can derive the effective halo mass using eq. 23 together with the dark matter halo mass function.

We compare the reported satellite fraction,  $f_s^{\text{ma}}$ , of Mandelbaum et al. (2006) and the derived effective halo mass,  $M_{\text{eff}}$ , with our results in Figure 10 and Figure 11. Figure 10 shows that the galaxies sampled by Mandelbaum et al. (2006) have a very similar effective halo mass compared to 6dFGS galaxies. The 6dFGS measurements follow the trend of early-type galaxies in these measurements, while the late types have a slightly smaller effective halo mass. However, the errors in case of the results by Mandelbaum et al. (2006) don't allow one to distinguish between the early and late type galaxies. Since effective halo mass and satellite fraction are strongly linked, we also compare the satellite fraction of Mandelbaum et al. (2006) with our results in Figure 11. Both figures also include the fit to sample  $S_2$  where we assumed  $\sigma_8 = 0.9$ , instead of the standard  $\sigma_8 = 0.8$ , since this agrees with the assumption in Mandelbaum et al. (2006). A larger  $\sigma_8$  increases the ef-



**Figure 10.** The effective dark matter halo mass of the four 6dFGS sub-samples (blue data points) compared to early type (black triangles) and late type (black squares) galaxies from Mandelbaum et al. (2006). In red we also include the fit to sample  $S_2$  which assumes  $\sigma_8 = 0.9$  in agreement with the assumptions of Mandelbaum et al. (2006), while the blue data points assume  $\sigma_8 = 0.8$ .



**Figure 11.** The satellite fraction of the four 6dFGS sub-samples (blue data points) compared to early type (black triangles) and late type (black squares) galaxies from Mandelbaum et al. (2006). In red we also include the fit to sample  $S_2$  which assumes  $\sigma_8 = 0.9$  in agreement with the assumptions of Mandelbaum et al. (2006), while the blue data points assume  $\sigma_8 = 0.8$ .

fective halo mass by a small amount and hence does not impact this comparison significantly.

Overall we see very good agreement between our results and the results of Mandelbaum et al. (2006), which is reassuring since the two techniques are subject to different systematic uncertainties. We can also emphasise that the 6dFGS results are extremely precise compared to the lensing results.

#### 8.4 Comparison to other studies

Meneux et al. (2007) study galaxy cluster mass dependence

on stellar mass in the VIMOS-VLT Deep Survey (VVDS) at redshift  $0.5 < z < 1.2$ . The stellar masses in this sample cover the range  $\log_{10}(M_{\text{stellar}}/h^{-2}M_{\odot}) = 8.7 - 10.7$ . To quantify the clustering, they use power laws fits to the projected correlation function  $w_p(r_p)$  and found an evolution in  $r_0$  from  $2.76 \pm 0.17 h^{-1}$  Mpc at  $\log_{10}(M_{\text{stellar}}/h^{-2}M_{\odot}) > 8.7$  to  $r_0 = 4.28 \pm 0.45 h^{-1}$  Mpc at  $\log_{10}(M_{\text{stellar}}/h^{-2}M_{\odot}) > 10.2$ . The slope changes over the same range from  $\gamma = 1.67 \pm 0.08$  to  $\gamma = 2.28 \pm 0.28$ . Comparing the results by Meneux et al. (2007) with Li et al. (2006) who used the clustering of 200 000 SDSS galaxies at  $z = 0.15$ , showed that the evolution of the amplitude and shape of the correlation function  $w_p(r_p)$  with redshift is faster for low stellar mass objects than for high stellar mass objects. At low stellar mass, the amplitude of  $w_p(r_p)$  increases by a factor  $\sim 2 - 3$  from high to low redshift, while at the high stellar mass range  $\log_{10}(M_{\text{stellar}}/h^{-2}M_{\odot}) = 10.2 - 10.7$ , the amplitude at  $z \sim 0.85$  and  $z \sim 0.15$  are very similar, within the error bars. In 6dFGS we found a significantly larger clustering amplitude, ranging from  $r_0 = 5.14 \pm 0.23 h^{-1}$  Mpc at  $\log_{10}(M_{\text{stellar}}/h^{-2}M_{\odot}) = 10.41$  to  $r_0 = 6.21 \pm 0.17 h^{-1}$  Mpc at  $\log_{10}(M_{\text{stellar}}/h^{-2}M_{\odot}) = 10.73$ , while our power law index  $\gamma$  stays constant at  $\gamma \approx 1.84$ . This shows that the trend observed in Meneux et al. (2007) and Li et al. (2006) continues to the 6dFGS redshifts, even for high stellar mass galaxies.

## 9 CONCLUSIONS

We present in this paper an analysis of the clustering properties of four stellar mass selected volume-limited sub-samples of galaxies from the 6dF Galaxy Survey. The stellar masses are calculated using the  $J$ -band magnitude and the  $b_J - r_F$  colour following the technique by Bell & De Jong (2001). The average log-stellar mass for the different sub-samples ranges from  $\log_{10}(M_{\text{stellar}}/h^{-2}M_{\odot}) = 10.41 - 10.73$ . Our analysis has the following main results:

- The projected correlation function,  $w_p(r_p)$ , for the two low stellar mass sub-samples ( $S_1$  and  $S_2$ ) can be described by a power law with an acceptable  $\chi^2$ , while the two high stellar mass sub-samples ( $S_3$  and  $S_4$ ) have a reduced  $\chi^2$  of 1.48 and 1.63, respectively. We also find patterns in the deviations between the best power law fit and the data, which can naturally be explained within the halo model. This is in agreement with theoretical studies (Watson et al. 2011), predicting that the disagreement of  $w_p(r_p)$  with a power law fit should grow with clustering amplitude.

- We used an HOD parameterisation with three free parameters ( $M_1$ ,  $\alpha$ ,  $M_{\text{min}}$ ), representing the typical halo mass, satellite power law index and minimum host halo mass, respectively. The minimum halo mass is fixed by the galaxy number density for all our parameter fits. We performed fits to the projected correlation functions of the four different volume-limited sub-samples. We tested alternative HOD parameterisations, but found that our data does not justify more free parameters.  $M_1$  and  $\alpha$  show increasing trends for increasing stellar mass, with  $\log_{10}(M_1/h^{-1}M_{\odot})$  ranging from 13.4 – 14 and  $\alpha$  from 1.21 – 1.4. This means that galaxies with larger stellar mass populate larger dark matter halos, which accrete satellites faster with increasing mass,

compared to dark matter halos populated by galaxies with lower stellar mass.

- From the halo model, we can derive averaged parameters for the four samples such as the satellite fraction, effective dark matter halo mass and the effective galaxy bias. We found that the satellite fraction decreases with stellar mass from 21% at  $\log_{10}(M_{\text{stellar}}/h^{-2} M_{\odot}) = 10.41$  to 12% at  $\log_{10}(M_{\text{stellar}}/h^{-2} M_{\odot}) = 10.73$ . The effective dark matter halo mass stays constant at  $\log_{10}(M_{\text{eff}}/h^{-1} M_{\odot}) \approx 13.55$  for all four sub-samples. The effective galaxy bias increases with stellar mass indicating that galaxies with higher stellar mass reside in denser environments. The increase in the effective galaxy bias can be described by  $(1.05 + M_{\text{stellar}}/M_*) \times (\sigma_8/0.8)$  with  $M_* = 1.18 \times 10^{11} h^{-2} M_{\odot}$ .

- We use two semi-analytic models, based on the Millennium Simulation (Croton et al. 2006 and Bower et al. 2006) to derive 6dFGS mock surveys. We compare the results of these semi-analytic models with our measurements. The Croton et al. (2006) model under-predicts the satellite fraction, while the Bower et al. (2006) model is in better agreement with our observations, although it slightly over-predicts the satellite fraction. Since the effective dark matter halo mass is strongly linked to the satellite fraction, the Bower et al. (2006) model prediction of  $M_{\text{eff}}$  is again in better agreement with our observations. 6dFGS allows a powerful test of semi-analytic models, because of (1) the robust stellar mass estimates and (2) the focus on “red satellites”, which semi-analytic models struggled to reproduce in the past. Our results can be used as a new constraint on semi-analytic models in the future.

- For the four volume-limited samples we find a constant scaling relation between  $M_1$  and  $M_{\text{min}}$  of  $M_1 \approx 22 M_{\text{min}}$ , in agreement with studies of SDSS galaxies (Zehavi et al. 2005b, 2010). This indicates that 6dFGS galaxies populate dark matter halos in a similar way to SDSS galaxies. However, we see a wide spectrum of this ratio ranging from  $> 40$  in Abbas et al. (2010) to  $< 10$  in (Matsuoka et al. 2011; Wake et al. 2011) indicating that this ratio is not universal for all types of galaxies.

- We compare our results with the results of Mandelbaum et al. (2006) from galaxy-galaxy weak lensing. We find overall good agreement which represents a valuable crosscheck for these two different clustering measurements. Although our analysis depends on slightly more assumptions and covers a smaller range in stellar mass, the 6dFGS results are extremely precise compared to the lensing results.

## ACKNOWLEDGEMENTS

The authors thank Alex Merson for providing the random mock generator. We also acknowledge Rachel Mandelbaum, Darren Croton, Danail Obreschkow, Alan Duffy, Martin Meyer and Michael Brown for helpful discussions. Thanks also to the referee for detailed feedback which helped to improve the final version of this paper.

F.B. is supported by the Australian Government through the International Postgraduate Research Scholarship (IPRS) and by scholarships from ICRAR and the AAO. The 6dF Galaxy Survey was funded in part by an Australian

Research Council Discovery–Projects Grant (DP-0208876), administered by the Australian National University.

The Millennium Run simulation used in this paper was carried out by the Virgo Supercomputing Consortium at the Computing Centre of the Max-Planck Society in Garching. The semi-analytic galaxy catalogue is publicly available at <http://www.mpa-garching.mpg.de/galform/agnpaper> and <http://galaxy-catalogue.dur.ac.uk:8080/Millennium/>.

## REFERENCES

- Abbas U. *et al.*, arXiv:1003.6129 [astro-ph.CO].
- Akaike H. IEEE Transactions on Automatic Control 19 (6): 716723
- Bell E. F. and de Jong R. S., Astrophys. J. **550**, 212 (2001) [astro-ph/0011493].
- Benson A. J., Lacey C. G., Baugh C. M., Cole S. and Frenk C. S., Mon. Not. Roy. Astron. Soc. **333**, 156 (2002) [astro-ph/0108217].
- Berlind A. A. and Weinberg D. H., Astrophys. J. **575** (2002) 587 [arXiv:astro-ph/0109001].
- Berlind A. A. *et al.*, Astrophys. J. **593**, 1 (2003) [astro-ph/0212357].
- Bertone S., De Lucia G. and Thomas P. A., Mon. Not. Roy. Astron. Soc. **379**, 1143 (2007) [astro-ph/0701407].
- Beutler F. *et al.*, Mon. Not. Roy. Astron. Soc. **416**, (2011) 3017B arXiv:1106.3366 [astro-ph.CO].
- Beutler F. *et al.*, Mon. Not. Roy. Astron. Soc. **423**, (2012) 3430B arXiv:1204.4725 [astro-ph.CO].
- Blake C., Collister A. and Lahav O., Mon. Not. Roy. Astron. Soc. **385**, 1257 (2008) [arXiv:0704.3377 [astro-ph]].
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S. and Lacey C. G., Mon. Not. Roy. Astron. Soc. **370**, 645 (2006) [astro-ph/0511338].
- Brown M. J. I. *et al.*, Astrophys. J. **682**, 937 (2008) [arXiv:0804.2293 [astro-ph]].
- Bruzual G. and Charlot S., Astrophys. J. **405**, 538 (1993).
- Bullock J. S. *et al.*, Mon. Not. Roy. Astron. Soc. **321** (2001) 559 [arXiv:astro-ph/9908159].
- Coil A. L. *et al.*, Astrophys. J. **638** (2006) 668 [astro-ph/0507647].
- Conroy C., Wechsler R. H. and Kravtsov A. V., Astrophys. J. **647**, 201 (2006) [astro-ph/0512234].
- Cooray A. and Sheth R. K., Phys. Rept. **372** (2002) 1 [arXiv:astro-ph/0206508].
- Croton D. J. *et al.*, Mon. Not. Roy. Astron. Soc. **365**, 11 (2006) [astro-ph/0508046].
- Davis M. and Peebles P. J. E., Astrophys. J. **267** (1982) 465.
- de la Torre S. *et al.*, arXiv:1010.0360 [astro-ph.CO].
- De Lucia G. and Blaizot J., Mon. Not. Roy. Astron. Soc. **375** (2007) 2 [astro-ph/0606519].
- Drory N., Bender R. and Hopp U., Astrophys. J. **616**, L103 (2004) [astro-ph/0410084].
- Duffy A. R., Schaye J., Kay S. T. and Dalla Vecchia C., Mon. Not. Roy. Astron. Soc. **390**, L64 (2008) [arXiv:0804.2486 [astro-ph]].
- Font A. S. *et al.*, Mon. Not. Roy. Astron. Soc. **389** (2008) 1619 [arXiv:0807.0001 [astro-ph]].
- Gallazzi A. and Bell E. F., Astrophys. J. Suppl. **185**, 253 (2009) [arXiv:0910.1591 [astro-ph.CO]].

- Grillo C., Gobat R., Rosati P. and Lombardi M., *A&A* **477**, L25 (2008) [arXiv:0712.0680 [astro-ph]].
- Guo Q., White S., Li C. and Boylan-Kolchin M., *Mon. Not. Roy. Astron. Soc.* **404**, 1111 (2010) [arXiv:0909.4305 [astro-ph.CO]].
- Hauser M. G., & Peebles, P. J. E., *ApJ*, 185, 757 (1973)
- Peebles, P. J. E., & Hauser, M. G., *ApJS*, 28, 19 (1974)
- Hawkins E. *et al.*, *Mon. Not. Roy. Astron. Soc.* **346**, 78 (2003) [arXiv:0212375].
- Helly J. C., Cole S., Frenk C. S., *et al.* 2003, *Mon. Not. Roy. Astron. Soc.* **338**, 903
- Jarrett T. H., Chester T., Cutri R., Schneider S., Skrutskie M. and Huchra J. P., *Astron. J.* **119** (2000) 2498 [arXiv:astro-ph/0004318].
- Jing Y. P., Mo H. J. and Borner G., *Astrophys. J.* **494**, 1 (1998) [astro-ph/9707106].
- Jones D. H. *et al.*, *MNRAS* **355** (2004) 747 [arXiv:astro-ph/0403501].
- Jones D. H., Saunders W., Read M. and Colless M., [astro-ph/0505068].
- Jones D. H., Peterson B. A., Colless M. and Saunders W., *MNRAS* **369** (2006) 25 [Erratum-ibid. **370** (2006) 1583] [arXiv:astro-ph/0603609].
- Jones D. H. *et al.*, arXiv:0903.5451 [astro-ph.CO].
- Kannappan S. J. and Gawiser E., *Astrophys. J.* **657** (2007) L5 [astro-ph/0701749 [ASTRO-PH]].
- Kauffmann G., White S. D. M. and Guiderdoni B., *Mon. Not. Roy. Astron. Soc.* **264**, 201 (1993).
- Kauffmann G. *et al.* [SDSS Collaboration], *Mon. Not. Roy. Astron. Soc.* **341** (2003) 33 [astro-ph/0204055].
- Kimm T. *et al.*, arXiv:0810.2794 [astro-ph].
- Kitzbichler M. G. and White S. D. M., *Mon. Not. Roy. Astron. Soc.* **376**, 2 (2007) [astro-ph/0609636].
- Komatsu E. *et al.* [WMAP Collaboration], *Astrophys. J. Suppl.* **192** (2011) 18 [arXiv:1001.4538 [astro-ph.CO]].
- Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottloeber S., Allgood B. and Primack J. R., *Astrophys. J.* **609**, 35 (2004) [astro-ph/0308519].
- Kroupa P., *Mon. Not. Roy. Astron. Soc.* **322**, 231 (2001) [astro-ph/0009005].
- Landy S. D. and Szalay A. S., *Astrophys. J.* **412** (1993) 64.
- Leauthaud A., Tinker J., Behroozi P. S., Busha M. T. and Wechsler R., *Astrophys. J.* **738**, 45 (2011) [arXiv:1103.2077 [astro-ph.CO]].
- Leauthaud A., *et al.*, *Astrophys. J.* **744**, 159 (2012) [arXiv:1104.0928 [astro-ph.CO]].
- Lee K., Giallisco K., Gnedin O. Y., Somerville R., Ferguson H., Dickinson M. and Ouchi M., *Astrophys. J.* **642** (2006) 63 [astro-ph/0508090].
- Lewis A. and Bridle S., *Phys. Rev. D* **66** (2002) 103511 [astro-ph/0205436].
- Li C., Kauffmann G., Jing Y. P. and White S. D. M., *Mon. Not. Roy. Astron. Soc.* **368** (2006) 21 [astro-ph/0509873].
- Liu L., Yang X., Mo H. J., van den Bosch F. C. and Springel V., *Astrophys. J.* **712** (2010) 734 [arXiv:0912.1257 [astro-ph.CO]].
- Longhetti M. and Saracco P., *Mon. Not. Roy. Astron. Soc.* **394** 774 (2009) arXiv:0811.4041 [astro-ph].
- Ma C. -P. and Fry J. N., *Astrophys. J.* **543**, 503 (2000) [astro-ph/0003343].
- Magliocchetti M. and Porciani C., *Mon. Not. Roy. Astron. Soc.* **346**, 186 (2003) [astro-ph/0304003].
- Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M. and Brinkmann J., *Mon. Not. Roy. Astron. Soc.* **368**, 715 (2006) [astro-ph/0511164].
- Matsuoka Y., Masaki S., Kawara K. and Sugiyama N., *Mon. Not. Roy. Astron. Soc.* **410**, 548 (2011) [arXiv:1008.0516 [astro-ph.CO]].
- Meneux B. *et al.*, [arXiv:0706.4371 [astro-ph]].
- More S., van den Bosch F. C., Cacciato M., Mo H., Yang X. and Li R., *Mon. Not. Roy. Astron. Soc.* **392**, 801 (2009) [arXiv:0807.4532 [astro-ph]].
- Navarro J. F., Frenk C. S. and White S. D. M., *Astrophys. J.* **490**, 493 (1997) [arXiv:astro-ph/9611107].
- Peacock J. A. and Smith R. E., *Mon. Not. Roy. Astron. Soc.* **318**, 1144 (2000) [astro-ph/0005010].
- Peebles P. J. E., *ApJ*, 185, 413 (1973)
- Peebles P. J. E., *AJ*, 189, L51 (1974)
- Porciani C., Magliocchetti M. and Norberg P., *Mon. Not. Roy. Astron. Soc.* **355**, 1010 (2004) [astro-ph/0406036].
- Scoccimarro R., Sheth R. K., Hui L., & Jain B., *ApJ*, 546, 20 (2001)
- Scoccimarro R., Sheth, R. K., Hui, L., & Jain, B. *ApJ*, 546, 20 (2001)
- Seljak U., *Mon. Not. Roy. Astron. Soc.* **318**, 203 (2000) [astro-ph/0001493].
- Simha V., Weinberg D. H., Dave R., Gnedin O. Y., Katz N. and Keres D., *Mon. Not. Roy. Astron. Soc.* **399**, 650 (2009) [arXiv:0809.2999 [astro-ph]].
- Smith R. E. *et al.* [The Virgo Consortium Collaboration], *MNRAS* **341** (2003) 1311 [arXiv:astro-ph/0207664].
- Snaith O. N., Gibson B. K., Brook C. B., Courty S., Sanchez-Blazquez P., Kawata D., Knebe A. and Sales L. V., arXiv:1104.2447 [astro-ph.GA].
- Spergel D. N. *et al.* [WMAP Collaboration], *Astrophys. J. Suppl.* **148** (2003) 175 [astro-ph/0302209].
- Springel V., White S. D. M., Tormen G. and Kauffmann G., *Mon. Not. Roy. Astron. Soc.* **328**, 726 (2001) [astro-ph/0012055].
- Springel V. *et al.*, *Nature* **435**, 629 (2005) [astro-ph/0504097].
- Tinker J. L., Weinberg D. H., Zheng Z. and Zehavi I., *Astrophys. J.* **631**, 41 (2005) [arXiv:astro-ph/0411777].
- Tinker J. L., Weinberg D. H. and Zheng Z., *Mon. Not. Roy. Astron. Soc.* **368**, 85 (2006) [arXiv:astro-ph/0501029].
- Tinker J. L., Norberg P., Weinberg D. H. and Warren M. S., *Astrophys. J.* **659**, 877 (2007) [astro-ph/0603543].
- Tinker J. L. *et al.*, *Astrophys. J.* **688**, 709 (2008) [arXiv:0803.2706 [astro-ph]].
- Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G. and Gottloeber S., arXiv:1001.3162 [astro-ph.CO].
- Totsuji, H., & Kihara, T., *PASJ*, 21, 221 (1969)
- van den Bosch F. C., Yang X. and Mo H. J., *Mon. Not. Roy. Astron. Soc.* **340** (2003) 771 [arXiv:astro-ph/0210495].
- Wake D. A., *et al.*, *Astrophys. J.* **728** (2011) 46 [arXiv:1012.1317 [astro-ph.CO]].
- Watson D. F., Berlind A. A. and Zentner A. R., *Astrophys. J.* **738**, 22 (2011) [arXiv:1101.5155 [astro-ph.CO]].
- Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., Croton D. J. and Moore B., *Mon. Not. Roy. Astron. Soc.* **372** (2006) 1161 [astro-ph/0606458].
- White S. D. M. and Frenk C. S., *Astrophys. J.* **379**, 52

- (1991).
- Worthey G., *Astrophys. J. Suppl.* **95** (1994) 107.
- Yang X. -H., Mo H. J., Jing Y. P. and van den Bosch F. C., *Mon. Not. Roy. Astron. Soc.* **358** (2005) 217 [astro-ph/0410114].
- Yang X., Mo H. J. and van den Bosch F. C., *Astrophys. J.* **676** (2008) 248 [arXiv:0710.5096 [astro-ph]].
- Zehavi I. *et al.* 2005a [SDSS Collaboration], *Astrophys. J.* **621** (2005) 22 [arXiv:astro-ph/0411557].
- Zehavi I. *et al.* 2005b [The SDSS Collaboration], *Astrophys. J.* **630**, 1 (2005) [astro-ph/0408569].
- Zehavi I. *et al.* [The SDSS Collaboration], *Astrophys. J.* **736**, 59 (2011) [arXiv:1005.2413 [astro-ph.CO]].
- Zheng Z. *et al.*, *Astrophys. J.* **633** (2005) 791 [arXiv:astro-ph/0408564].
- Zheng Z., Coil A. L. and Zehavi I., *Astrophys. J.* **667** (2007) 760 [astro-ph/0703457 [ASTRO-PH]].
- Zheng Z., Zehavi I., Eisenstein D. J., Weinberg D. H. and Jing Y., *Astrophys. J.* **707**, 554 (2009) [arXiv:0809.1868 [astro-ph]].

## APPENDIX A: NORMALISATION CONSTANT IN Mandelbaum et al. 2006

We can determine the normalisation constant  $k$  of eq. 27 by matching to the satellite fraction  $f_s^{\text{ma}}$ , which in terms of the mass function  $dn(M)/dM$  is given by

$$f_s^{\text{ma}} = \frac{\int_0^\infty dM \frac{dn(M)}{dM} N_s(M)}{\int_0^\infty dM \frac{dn(M)}{dM} [N_c(M) + N_s(M)]} \quad (\text{A1})$$

$$= \frac{kH}{\int_{M_{\min}}^\infty dM \frac{dn(M)}{dM} + kH} \quad (\text{A2})$$

with

$$H = \frac{1}{3M_{\min}} \int_{M_{\min}}^{3M_{\min}} dM \frac{dn(M)}{dM} M^2 + \int_{3M_{\min}}^\infty dM \frac{dn(M)}{dM} M, \quad (\text{A3})$$

which leads to

$$k = \frac{f_s^{\text{ma}} \int_{M_{\min}}^\infty dM \frac{dn(M)}{dM}}{(1 - f_s^{\text{ma}})H}, \quad (\text{A4})$$

where  $k$  has units of  $1/h^{-1}M_\odot$  and is constant for a given set of  $f_s^{\text{ma}}$  and  $M_{\min}$ .