The abundance and radial distribution of satellite galaxies

Frank C. van den Bosch,[1]* Xiaohu Yang,[2] H. J. Mo[2] and Peder Norberg[1]

[1]*Department of Physics, Swiss Federal Institute of Technology, ETH Hönggerberg, CH-8093, Zurich, Switzerland*
[2]*Department of Astronomy, University of Massachussets, 710 North Pleasant Street, Amherst MA 01003-9305, USA*

ABSTRACT

Using detailed mock galaxy redshift surveys (MGRSs) we investigate the abundance and radial distribution of satellite galaxies. The mock surveys are constructed using large numerical simulations and the conditional luminosity function (CLF), and are compared against data from the Two Degree Field Galaxy Redshift Survey (2dFGRS). We use Monte Carlo Markov chains to explore the full posterior distribution of the CLF parameter space, and show that the average relation between light and mass is tightly constrained and in excellent agreement with our previous models and with that of Vale & Ostriker. The radial number density distribution of satellite galaxies in the 2dFGRS reveals a pronounced absence of satellites at small projected separations from their host galaxies. This is (at least partly) owing to the overlap and merging of galaxy images in the 2dFGRS parent catalogue. Owing to the resulting close-pair incompleteness we are unfortunately unable to put meaningful constraints on the radial distribution of satellite galaxies; the data are consistent with a radial number density distribution that follows that of the dark matter particles, but we cannot rule out alternatives with a constant number density core. Marginalizing over the full CLF parameter space, we show that in a ΛCDM concordance cosmology the observed abundances of host and satellite galaxies in the 2dFGRS indicate a power spectrum normalization of $\sigma_8 \simeq 0.7$. The same cosmology but with $\sigma_8 = 0.9$ is unable to match simultaneously the abundances of host and satellite galaxies. This confirms our previous conclusions based on the pairwise peculiar velocity dispersions and the group multiplicity function.

Key words: methods: statistical – galaxies: formation – galaxies: fundamental parameters – galaxies: haloes – cosmological parameters – dark matter.

1 INTRODUCTION

In the hierarchical formation scenario, satellite galaxies are associated with dark matter subhaloes, which, some time in the past, were accreted by their current host (or parent) halo. As satellites orbit their host galaxies they are subjected to various forces that try to dissolve them: dynamical friction, tides from the central object(s) and impulsive collisions with other satellites. A detailed understanding of the abundance and radial distribution of satellite systems, therefore, provides important constraints on the outcome of these various physical processes which are an essential ingredient of galaxy formation.

Traditionally, satellite galaxies have mainly been used as kinematic tracers of the dark matter potential well of the parent halo (e.g. Little & Tremaine 1987; Zaritsky et al. 1993; Carlberg et al. 1996; Zaritsky et al. 1997; Evans et al. 2000; McKay et al. 2002; Prada et al. 2003; Brainerd & Specian 2003; van den Bosch et al. 2004b). As satellite galaxies are distributed over the entire halo,

they are ideally suited to measure the total virial mass, more so than for example rotation curves, which only probe the potential out to a fraction of the virial radius. The *distribution* of satellite galaxies, however, has received considerably less attention. In addition to a few studies aimed at testing the claim by Holmberg (1969) that the azimuthal distribution of satellite galaxies is anisotropic with respect to the orientation of the host galaxy, the so-called Holmberg effect (Zaritsky et al. 1997; Zaritsky & Gonzales 1999; Sales & Lambas 2004), relatively few studies have focused on the *radial* distribution of satellite galaxies (but see Lake & Tremaine 1980; Vader & Sandage 1991; Lorrimer et al. 1994; Willman et al. 2004).

The interest in satellite galaxies has recently increased considerably, largely owing to the dramatic increase in computing power that has made it possible to resolve dark matter subhaloes in cosmological numerical simulations (Tormen 1997; Ghigna et al. 1998; Avila-Reese et al. 1999; Klypin et al. 1999; Moore et al. 1999; Stoehr et al. 2002; Kravtsov et al. 2004a; De Lucia et al. 2004; Diemand, Moore & Stadel 2004; Gill et al. 2004a; Gill, Knebe & Gibson 2004b; Weller, Ostriker & Bode 2004; Reed et al. 2004). This has resulted in a number of detailed studies of the abundances

*E-mail: vdbosch@phys.ethz.ch

and both the spatial and the velocity distribution of subhaloes. As satellite galaxies are thought to be associated with these subhaloes, this has opened the possibility to compare the statistical properties of satellite galaxies directly with those of dark matter subhaloes, resulting in two apparent inconsistencies.

First of all, the radial distribution of dark matter subhaloes is found to be spatially antibiased with respect to dark matter. However, the observed distribution of cluster galaxies seems to accurately follow a Navarro, Frenk & White (1997, hereafter NFW) distribution (Beers & Tonry 1986; Carlberg, Yee & Ellingson 1997; van der Marel et al. 2000; Lin, Mohr & Stanford 2004). Although this might signal an inconsistency of the standard CDM framework, semi-analytical models of galaxy formation that take the evolution of dark matter subhaloes into account can reproduce the observed distribution of cluster galaxies (Springel et al. 2001; Diaferio et al. 2001; Gao et al. 2004a; Kravtsov, Gnedin & Klypin 2004b). Part of the reason is that galaxies are dense concentrations of baryons in the centers of dark matter (sub)haloes, which are more resilient to tidal disruption than (the outer parts of) their dark matter haloes, but other processes, such as tidally induced star formation and tidal heating also play important roles (see e.g. Kravtsov et al. 2004b).

The second apparent inconsistency concerns the abundances of satellite galaxies. Although the subhalo mass function, when normalized to the mass of the parent halo, is found to be virtually independent of halo mass (Moore et al. 1999; De Lucia et al. 2004; Diemand et al. 2004; Weller et al. 2004; but see Gao et al. 2004b), the satellite populations of galaxy clusters are very different from those of galaxy-sized haloes (Klypin et al. 1999; Moore et al. 1999; D'onghia & Lake 2004). This must be telling us something important about the physics of galaxy formation, and numerous studies have focused on various mechanisms to explain this discrepancy between the number of predicted subhaloes and observed satellites (e.g. Kauffmann, White & Guiderdoni 1993; Bullock, Kravtsov & Weinberg 2000; Benson et al. 2002).

At this point in time there is a strong need for better observational constraints regarding the statistical properties of satellite galaxies. In particular, does the radial distribution of satellite galaxies in galaxy-sized haloes follow a NFW profile as for clusters, or does it reveal a constant density core, as found for dark matter subhaloes? How does the abundance of satellite galaxies depend on halo mass, or on the luminosity of the host galaxy? Is this mass-dependent abundance consistent with the full galaxy luminosity function and the halo mass function? The conditional luminosity function (hereafter CLF) formalism developed by Yang, Mo & van den Bosch (2003) and van den Bosch, Yang & Mo (2003a) is ideally suited to address these questions. It describes how many galaxies of given luminosity reside in a halo of given mass. In this paper we therefore use the CLF, constrained using the abundances and clustering properties of galaxies in the Two-Degree Field Galaxy Redshift Survey (2dFGRS, Colless et al. 2001), to study the abundances and radial distribution of satellite galaxies. We construct detailed mock galaxy redshift surveys (MGRSs), based on the CLF, for direct comparison with the 2dFGRS data. We show that a close-pair incompleteness in the 2dFGRS prevents us from putting significant constraints on the radial number density distribution of satellite galaxies, and that matching the abundances of host and satellite galaxies in the 2dF-GRS requires a ΛCDM concordance cosmology with a relatively low value of the power-spectrum normalization parameter σ_8.

This paper is organized as follows. In Section 2 we describe the CLF, and use a Monte Carlo Markov chain (hereafter MCMC) to fully sample the posterior distribution of the CLF. In Section 3 we describe how to use the CLF to construct detailed MGRSs for direct comparison with the 2dFGRS. The selection criteria for host and satellite galaxies are described in Section 4. Section 5 compares the projected, radial distribution of satellite galaxies extracted from the 2dFGRS with those obtained from our MGRSs, including a detailed discussion of incompleteness effects in the 2dFGRS. Section 6 compares the abundances of host and satellite galaxies between 2dFGRS and MGRS. We summarize our results in Section 7.

2 THE CONDITIONAL LUMINOSITY FUNCTION

Yang et al. (2003) and van den Bosch et al. (2003a) presented a new method to link the distribution of galaxies to that of dark matter haloes. This method is based on modelling the CLF, $\Phi(L|M)\,dL$, which gives the average number of galaxies with luminosity $L \pm dL/2$ that reside in a halo of mass M. This CLF is the direct link between the halo mass function $n(M)\,dM$, specifying the comoving number density of haloes of mass M, and the galaxy luminosity function $\Phi(L)\,dL$, specifying the comoving number density of galaxies with luminosity L, through

$$\Phi(L) = \int_0^\infty \Phi(L|M)\,n(M)\,dM. \tag{1}$$

In CDM cosmologies, more massive haloes are more strongly clustered (Cole & Kaiser 1989; Mo & White 1996, 2002). This means that information on the clustering strength of galaxies (of a given luminosity) contains information about the characteristic mass of the haloes in which they reside. Therefore, an observed luminosity function $\Phi(L)$ combined with measurements of the galaxy–galaxy two-point correlation function *as function of luminosity*, $\xi_{gg}(r, L)$, puts stringent constraints on $\Phi(L|M)$ (see Yang et al. 2003).

For a given CLF the luminosity function $\Phi(L)$ follows directly from equation (1) while, at sufficiently large separations r, the two-point correlation function is given by

$$\xi_{gg}(r, L) = b^2(L)\,\xi_{dm}(r). \tag{2}$$

Here $\xi_{dm}(r)$ is the dark matter mass correlation function and $b(L)$ is the average *bias* of galaxies of luminosity L, which derives from the CLF according to

$$b(L) = \frac{1}{\Phi(L)} \int_0^\infty \Phi(L|M)\,b(M)\,n(M)\,dM, \tag{3}$$

with $b(M)$ the bias of dark matter haloes of mass M.

Throughout this paper we compute the halo mass function using the form suggested by Sheth, Mo & Tormen (2001), which has been shown to be in excellent agreement with numerical simulations as long as halo masses are defined as the masses inside a sphere with an average overdensity of about 180 (Jing 1998; Sheth & Tormen 1999; Jenkins et al. 2001; White 2002). Therefore, in what follows we consistently use that definition of halo mass when referring to M. The halo bias, $b(M)$, is computed using the fitting formula of Seljak & Warren (2004). The linear power spectrum of density perturbations is computed using the transfer function of Eisenstein & Hu (1998), which properly accounts for the baryons, while the evolved, non-linear power spectrum, required to compute the dark matter correlation function, is computed using the fitting formula of Smith et al. (2003).

Throughout this paper we assume a flat ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $h = H_0/(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}) = 0.7$ and with a scale-invariant initial power spectrum. Unless stated otherwise, we adopt a normalization of $\sigma_8 = 0.9$.

2.1 Parametrization

The CLF is parametrized by a Schechter function

$$\Phi(L|M)\mathrm{d}L = \frac{\tilde{\Phi}^*}{\tilde{L}^*}\left(\frac{L}{\tilde{L}^*}\right)^{\tilde{\alpha}}\exp(-L/\tilde{L}^*)\,\mathrm{d}L, \tag{4}$$

where $\tilde{L}^* = \tilde{L}^*(M)$, $\tilde{\alpha} = \tilde{\alpha}(M)$ and $\tilde{\Phi}^* = \tilde{\Phi}^*(M)$ are all functions of halo mass M. We write the average, total mass-to-light ratio of a halo of mass M as

$$\langle M/L\rangle_M = \frac{1}{2}\left(\frac{M}{L}\right)_0\left[\left(\frac{M}{M_1}\right)^{-\gamma_1} + \left(\frac{M}{M_1}\right)^{\gamma_2}\right]. \tag{5}$$

This parametrization has four free parameters: a characteristic mass M_1, for which the mass-to-light ratio is equal to $(M/L)_0$, and two slopes, γ_1 and γ_2, that specify the behavior of $\langle M/L\rangle$ at the low- and high-mass ends, respectively. Motivated by observations (Bahcall, Lubin & Dorman 1995; Bahcall et al. 2000; Sanderson & Ponman 2003), which indicate a flattening of $\langle M/L\rangle_M$ on the scale of galaxy clusters, we set $\langle M/L\rangle_M = (M/L)_{\mathrm{cl}}$ for haloes with $M \geqslant 10^{14}\,h^{-1}\,\mathrm{M}_\odot$.

A similar parametrization is used for the characteristic luminosity $\tilde{L}^*(M)$

$$\frac{M}{\tilde{L}^*(M)} = \frac{1}{2}\left(\frac{M}{L}\right)_0 f(\tilde{\alpha})\left[\left(\frac{M}{M_1}\right)^{-\gamma_1} + \left(\frac{M}{M_2}\right)^{\gamma_3}\right], \tag{6}$$

with

$$f(\tilde{\alpha}) = \frac{\Gamma(\tilde{\alpha}+2)}{\Gamma(\tilde{\alpha}+1,1)}. \tag{7}$$

Here $\Gamma(x)$ is the Gamma function and $\Gamma(a,x)$ the incomplete Gamma function. This parametrization has two additional free parameters: a characteristic mass M_2 and a power-law slope γ_3. For $\tilde{\alpha}(M)$ we adopt a simple linear function of $\log(M)$

$$\tilde{\alpha}(M) = \alpha_{15} + \eta\log(M_{15}), \tag{8}$$

with M_{15} the halo mass in units of $10^{15}\,h^{-1}\,\mathrm{M}_\odot$, $\alpha_{15} = \tilde{\alpha}(M_{15}=1)$, and η describes the change of the faint-end slope $\tilde{\alpha}$ with halo mass. Note that once $\tilde{\alpha}$ and $\tilde{L}^*$ are given, the normalization $\tilde{\Phi}^*$ of the CLF is obtained through equation (5), using the fact that the total (average) luminosity in a halo of mass M is given by

$$\langle L\rangle_M = \int_0^\infty \Phi(L|M)\,L\,\mathrm{d}L = \tilde{\Phi}^*\,\tilde{L}^*\,\Gamma(\tilde{\alpha}+2). \tag{9}$$

Finally, we introduce the mass scale $M_{\min}$ below which we set the CLF to zero; i.e. we assume that no stars form inside haloes with $M < M_{\min}$. Motivated by reionization considerations (see Yang et al. 2003, for details) we adopt $M_{\min} = 10^9\,h^{-1}\,\mathrm{M}_\odot$ throughout.

2.2 Parameter fitting

The CLF, as specified above, has a total of eight free parameters: two characteristic masses; M_1 and M_2, three parameters that describe the various mass dependencies γ_1, γ_3 and η, two normalizations for the mass-to-light ratio, $(M/L)_0$ and $(M/L)_{\mathrm{cl}}$, and a normalization of the faint-end slope, α_{15}. Note that γ_2 is not a free parameter as it derives from requiring continuity in $\langle M/L\rangle_M$ across $M = 10^{14}\,h^{-1}\,\mathrm{M}_\odot$. The data that we use to constrain the CLF consists of the 2dFGRS luminosity function of Madgwick et al. (2002) and the galaxy–galaxy correlation lengths as a function of luminosity obtained from the 2dFGRS by Norberg et al. (2002a). In Yang et al. (2003) and van den Bosch et al. (2003a) we presented a number of CLFs that accurately fit these data and that were based on different

assumptions regarding the free parameters. Motivated by a number of independent observational constraints, the majority of these models were constrained to have $(M/L)_{\mathrm{cl}} \simeq 500\,h\,(\mathrm{M}/L)_\odot$. However, subsequent studies have shown that these CLF models predict too many rich galaxy groups (Yang et al. 2004b), and pairwise peculiar velocity dispersions that are too high (Yang et al. 2004a). Both these problems are alleviated by adopting a much higher cluster mass-to-light ratio of $(M/L)_{\mathrm{cl}} \simeq 900\,h\,(\mathrm{M}/L)_\odot$.

In this paper we adopt a different approach. Rather than fixing $(M/L)_{\mathrm{cl}}$ at a preferred value and using a minimization routine to search our multidimensional parameter space for the best-fitting parameters, we follow Yan, Madgwick & White (2003) and use a MCMC to fully describe the likelihood function in our multidimensional parameter space. This will allow us to more accurately investigate the freedom in cluster mass-to-light ratios. Readers not familiar with, or interested in, MCMCs are referred to Gamerman (1997) for details.

We start our MCMC from model D in van den Bosch, Yang & Mo (2004a) and allow a 'burn-in' of 10 000 random walk steps for the chain to equilibrate in the likelihood space. At any point in the chain we generate a new trial model by drawing the shifts in its eight free parameters from eight independent Gaussian distributions. The probability of accepting the trial model is

$$P_{\mathrm{accept}} = \begin{cases} 1.0 & \text{if } \chi^2_{\mathrm{new}} < \chi^2_{\mathrm{old}} \\ \exp\left[-\left(\chi^2_{\mathrm{new}} - \chi^2_{\mathrm{old}}\right)/2\right] & \text{if } \chi^2_{\mathrm{new}} \geqslant \chi^2_{\mathrm{old}} \end{cases}. \tag{10}$$

Here χ^2 is defined as $\chi^2 = \chi^2_\Phi + \chi^2_{r_0}$ with

$$\chi^2_\Phi = \sum_{i=1}^{N_\Phi}\left[\frac{\Phi(L_i) - \hat{\Phi}(L_i)}{\Delta\hat{\Phi}(L_i)}\right]^2 \tag{11}$$

and

$$\chi^2_{r_0} = \sum_{i=1}^{N_r}\left[\frac{r_0(L_i) - \hat{r}_0(L_i)}{\Delta\hat{r}_0(L_i)}\right]^2. \tag{12}$$

Here $\hat{\Phi}$ and $\hat{r}_0$ are the observed quantities and $N_\Phi = 35$ and $N_r = 8$ are the number of data points for the LF and the correlation lengths, respectively.

We construct a MCMC of 40 million steps, with an average acceptance rate of $\sim$12 per cent. To suppress the correlation power between neighbouring models in the chain, we thin the chain by a factor of 20 000. This results in a final MCMC consisting of 2000 independent models that properly sample the full posterior distribution. The hatched histograms in Fig. 1 plot the resulting distributions of parameters, with the best-fitting values indicated by a vertical line. The median and 68 per cent confidence intervals of these distributions are listed in Table 1 (model $\Lambda_{0.9}$). Note that these parameters are similar to those of models A–D presented in van den Bosch et al. (2004a).[1]

In Fig. 2 we show scatter plots for some of the most tightly correlated pairs of parameters. The number density of points in these plots is directly proportional to the probability density. The solid circles correspond to 12 specific models to be discussed in more detail below. The most pronounced correlations are between $(M/L)_{\mathrm{cl}}$, $(M/L)_0$ and γ_2: increasing the cluster mass-to-light ratio results in a decrease of the minimum mass-to-light ratio $(M/L)_0$. This simply

---

[1] Small differences result from the fact that here we use the halo bias fitting function of Seljak & Warren (2004), rather than that of Sheth et al. (2001), and the transfer function of Eisenstein & Hu (1998), rather than that of Efstathiou, Bond & White (1992).
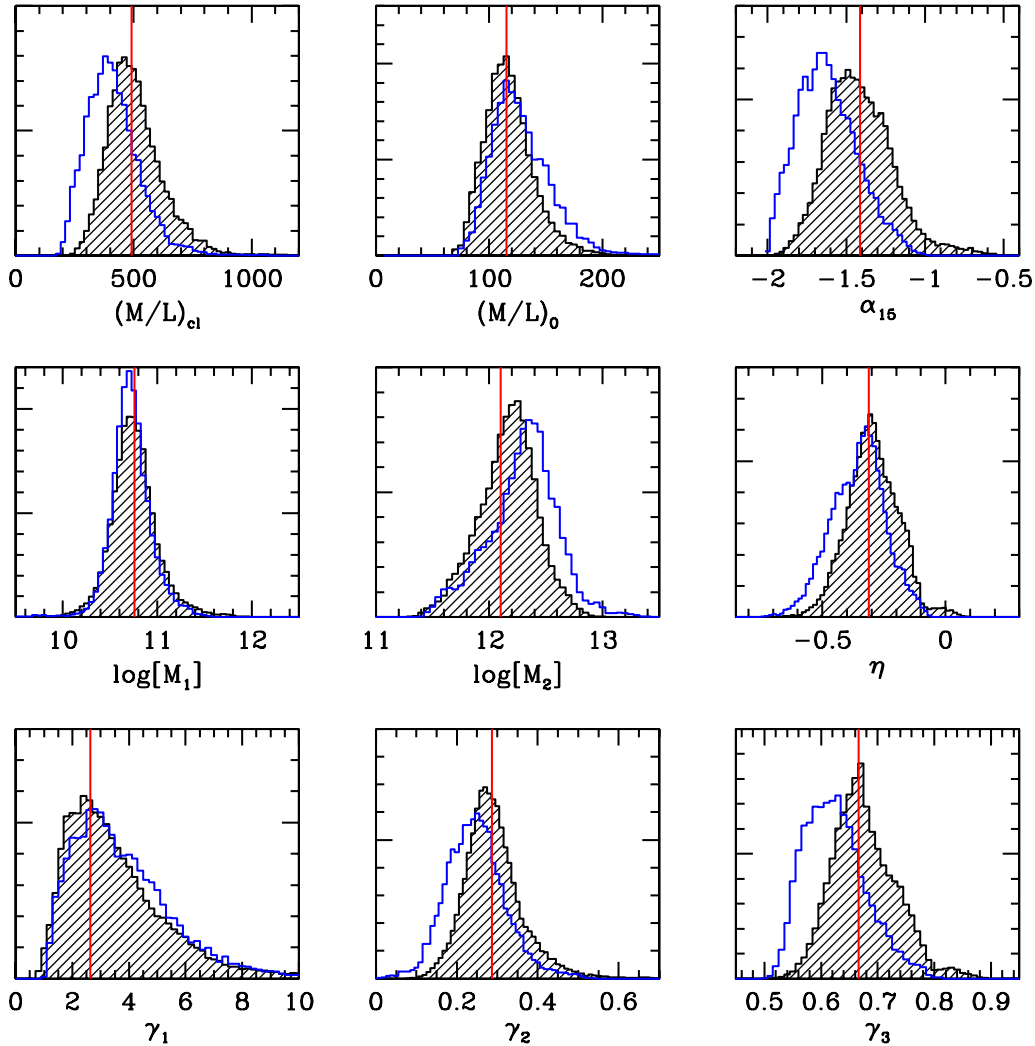
Figure 1. Constraints on the nine CLF parameters obtained from MCMCs with 2000 independent samples. The hatched (non-hatched) histograms correspond to ΛCDM concordance cosmologies with $\sigma_8 = 0.9$ (0.7), respectively. Vertical lines indicate the best-fitting parameters for the $\sigma_8 = 0.9$ cosmology. The median and 68 per cent confidence intervals of these distributions are listed in Table 1. Masses and mass-to-light ratios are in units of $h^{-1}\,M_\odot$ and $h\,(M/L)_\odot$, respectively.

reflects the conservation of the number of galaxies (constrained by the LF): putting fewer galaxies in clusters requires an increase in the occupation number in lower mass haloes. As γ_2 reflects the slope of $\langle M/L \rangle_M$ in between M_1 and the cluster scale, it is clear that γ_2 has to increase with increasing $(M/L)_{cl}$ and/or decreasing $(M/L)_0$. Another important correlation is between M_1 and γ_1, which is a direct consequence of the shallow faint-end slope of the galaxy LF. Owing to the much steeper low-mass end of the halo mass function, $\langle M/L \rangle_M$ has to increase with decreasing mass. Lowering M_1 results in a steeper increase, and therefore a larger value of γ_1.

Fig. 3 shows confidence levels on various quantities computed from our MCMC. The open circles with error bars in the upper left and upper middle panels indicate the data used to constrain the models. The shaded areas indicate the 68 and 99 per cent confidence levels on $\Phi(L)$ and $r_0(L)$ computed from the MCMC. Note the good agreement with the data, indicating that the CLF can accurately match the observed abundances and clustering properties of galaxies in the 2dFGRS. We emphasize that this is not a trivial result, as the data can only be fitted for a certain combination of cosmological parameters (van den Bosch, Mo & Yang 2003b).

The upper right-hand panel of Fig. 3 plots the faint-end slope of the CLF as a function of halo mass. The data clearly favors models in which $\tilde{\alpha}$ increases with decreasing halo mass. At around the cluster scale the models favor fairly steep faint-end slopes with $-1.0 \lesssim \tilde{\alpha} \lesssim -1.5$, which is in excellent agreement with independent studies of the luminosity functions in a number of individual clusters (e.g. Sandage, Bingelli & Tammann 1985; Beijersbergen et al. 2002; Trentham & Hodgkin 2002; Trentham & Tully 2002). The open circles with error bars indicate $\tilde{\alpha}(M)$ for three different halo masses M obtained by Eke et al. (2004) from an analysis of groups in the 2dFGRS, and are in good agreement with our model.

The lower left-hand panel of Fig. 3 plots the relation between halo mass M and the total halo luminosity L, the expectation value of which follows from the CLF according to equation (9). Note that the confidence levels are extremely tight, especially for the more massive haloes. The $L(M)$ relation reveals a dramatic break at around $M = M_1 \simeq 7 \times 10^{10}\,h^{-1}\,M_\odot$, with $L \propto M^{3.5 \pm 1.5}$ for $M \ll M_1$ and $L \propto M^{0.73 \pm 0.08}$ for $M \gg M_1$ (error bars are obtained from the confidence levels on γ_1 and γ_2, respectively). In a recent study, Vale & Ostriker (2004), under the assumption of

Table 1. CLF parameters.

ID (1)	$(M/L)_{cl}$ (2)	$(M/L)_0$ (3)	$\log M_1$ (4)	$\log M_2$ (5)	γ_1 (6)	γ_2 (7)	γ_3 (8)	α_{15} (9)	η (10)	χ^2 (11)
1	311.2	167.1	10.76	12.45	3.803	0.176	0.609	-1.71	-0.374	70.53
2	353.2	155.2	10.61	12.46	4.310	0.194	0.650	-1.62	-0.314	68.46
3	393.6	135.8	10.77	12.22	2.770	0.236	0.624	-1.60	-0.373	65.97
4	441.6	125.0	10.72	12.24	3.141	0.259	0.646	-1.55	-0.330	65.09
5	490.9	115.1	10.76	12.10	2.639	0.287	0.666	-1.41	-0.312	64.34
6	508.2	111.9	10.77	12.08	2.594	0.297	0.654	-1.46	-0.333	64.42
7	563.6	102.1	10.88	11.93	2.106	0.334	0.664	-1.31	-0.327	65.18
8	632.4	95.1	10.93	11.86	1.970	0.366	0.665	-1.29	-0.342	66.20
9	726.1	85.7	11.15	11.66	1.451	0.431	0.653	-1.23	-0.400	67.95
10	798.0	82.2	11.21	11.58	1.331	0.461	0.639	-1.25	-0.445	68.91
11	899.5	76.4	11.30	11.64	1.444	0.507	0.684	-1.10	-0.310	72.01
12	1106.6	71.4	11.71	11.48	0.976	0.652	0.656	-1.11	-0.401	72.99
$\Lambda_{0.9}$	500^{+130}_{-100}	115^{+23}_{-20}	$10.76^{+0.25}_{-0.21}$	$12.15^{+0.24}_{-0.31}$	$2.92^{+2.11}_{-1.15}$	$0.29^{+0.08}_{-0.06}$	$0.66^{+0.06}_{-0.04}$	$-1.44^{+0.21}_{-0.18}$	$-0.31^{+0.10}_{-0.10}$	–
$\Lambda_{0.7}$	400^{+120}_{-100}	125^{+30}_{-21}	$10.71^{+0.21}_{-0.17}$	$12.32^{+0.26}_{-0.35}$	$3.46^{+2.18}_{-1.35}$	$0.24^{+0.08}_{-0.07}$	$0.62^{+0.06}_{-0.05}$	$-1.64^{+0.21}_{-0.17}$	$-0.35^{+0.10}_{-0.12}$	–
$\Lambda^{ab}_{0.7}$	570^{+40}_{-40}	98^{+3}_{-4}	$10.86^{+0.04}_{-0.04}$	$11.96^{+0.10}_{-0.11}$	$2.45^{+0.55}_{-0.37}$	$0.34^{+0.02}_{-0.01}$	$0.63^{+0.03}_{-0.05}$	$-1.35^{+0.04}_{-0.25}$	$-0.35^{+0.08}_{-0.11}$	–

Parameters of CLF models. Column (1) lists the ID by which we refer to each CLF in the text. Columns (2)–(10) list the model parameters, and column (11) the value of $\chi^2 = \chi^2_\Phi + \chi^2_{r_0}$. Masses and mass-to-light ratios are in $h^{-1}\,M_\odot$ and $h\,(M/L)_\odot$, respectively. The first 12 lines (IDs 1–12) correspond to the best-fitting models extracted from the $\sigma_8 = 0.9$ MCMC for different bins in $(M/L)_{cl}$. They are shown as thick solid dots in Fig. 2. The last three lines list the median and 68 per cent confidence levels of the parameter probability distributions obtained from the MCMCs. Models $\Lambda_{0.9}$ and $\Lambda_{0.7}$ correspond to the MCMC of the ΛCDM cosmologies with $\sigma_8 = 0.9$ and $\sigma_8 = 0.7$, respectively. Model $\Lambda^{ab}_{0.7}$ is the same as Model $\Lambda_{0.7}$, except that we have weighted the MCMC samples by $\exp(-\chi^2_{ab}/2)$ (see Section 6.1).

a monotonic $L(M)$, obtained the relation between light and mass from a comparison of the galaxy luminosity function with the *total* halo mass function (counting both parent and subhaloes). They find that $L(M)$ changes from $L \propto M^4$ to $L \propto M^{0.9}$, in good agreement with our results. This is also evident from a direct comparison of their results (solid line in the lower left-hand panel of Fig. 3) with ours, which shows almost perfect agreement. It is extremely reassuring that two such wildly different methods yield results in such good agreement. This combined with the extremely tight confidence levels obtained from our CLF analysis suggests that we have established a remarkably robust connection between galaxy light and halo mass (at least for the concordance cosmology adopted here).

The lower, middle panel of Fig. 3 plots the corresponding mass-to-light ratios as function of halo mass. The pronounced minimum in M/L indicates that galaxy formation is most efficient in haloes with masses in the range $5 \times 10^{10}\,h^{-1}\,M_\odot \lesssim M \lesssim 10^{12}\,h^{-1}\,M_\odot$. For less massive haloes, the mass-to-light ratio increases drastically with decreasing halo mass, which is required to bring the steep slope of the halo mass function at low M in agreement with the relatively shallow faint-end slope of the observed LF. It indicates that galaxy formation needs to become extremely inefficient in haloes with $M \lesssim 5 \times 10^{10}\,h^{-1}\,M_\odot$ to prevent an overabundance of faint galaxies. The increase in $\langle M/L \rangle_M$ from $M \sim 10^{11}\,h^{-1}\,M_\odot$ to $M \sim 10^{14}\,h^{-1}\,M_\odot$ is associated with the decreasing ability of the gas to cool with increasing halo mass (e.g. White & Rees 1978; van den Bosch 2002). Finally, the sharp turn-over to a constant $\langle M/L \rangle_M$ for haloes with $M \geqslant 10^{14}\,h^{-1}\,M_\odot$ is a direct reflection of our CLF parametrization (see Section 2.1). The various lines correspond to models A–D from van den Bosch et al. (2003a), and are in excellent agreement with the confidence intervals obtained here. This indicates that the CLF models used in our previous work (Yang et al. 2004a,b; van den Bosch et al. 2003a,b, 2004a,b; Mo et al. 2004; Wang et al. 2004) are perfectly consistent with the parameter constraints obtained here using the MCMC.

Finally, the lower right-hand panel of Fig. 3 plots predictions for the number of galaxies (with $M_{b_J} - 5 \log h < -14$) as a function of halo mass. These are derived from the CLF according to

$$\langle N \rangle_M = \int_{L_{min}}^{\infty} \Phi(L|M)\,dL, \tag{13}$$

with L_{min} being the minimum luminosity considered ($L_{min} = 5.2 \times 10^7\,h^{-2}\,L_\odot$ in our case). Although the exact shape and normalization of $\langle N \rangle_M$ depends on L_{min} (e.g. van den Bosch et al. 2003a), the power-law behavior combined with a shoulder plus break at low M values is in good agreement with a number of studies based on halo occupation numbers (Seljak 2000; Scranton 2003; Berlind et al. 2003; Magliocchetti & Porciani 2003; Kravtsov et al. 2004a).

We can compare these predictions with the number of galaxies with $M_{b_J} - 5 \log h < -14$ in the MW and M31 systems. Using the data collected by Mateo (1998), and adopting $h = 0.7$, we find $N = 3$ for the MW (MW, LMC and SMC), and $N = 6$ for M31 (M31, M32, M33, NGC 205, NGC 147 and IC 10). Adopting the virial masses for the MW ($7 \times 10^{11}\,h^{-1}\,M_\odot$) and M31 ($1.1 \times 10^{12}\,h^{-1}\,M_\odot$) obtained by Klypin, Zhao & Somerville (2002), and converting these to our definition of halo mass M (see Section 2), we obtain the results indicated by the solid dot and the asterisk, respectively. Although these occupation numbers lie somewhat above our predictions, we caution that the latter indicate the 68 and 99 per cent confidence levels on the *average* $\langle N \rangle_M$; they do not indicate the scatter. Furthermore, the virial masses are quite uncertain; if the true virial masses of the MW and M31 are twice as large as the values advocated by Klypin et al. (see, for instance Hernández, Avila-Reese & Firmani 2001), they nicely match our CLF predictions.

3 MOCK GALAXY REDSHIFT SURVEYS

To properly interpret the 2dFGRS data on the abundance and radial distribution of satellite galaxies we construct detailed mock galaxy
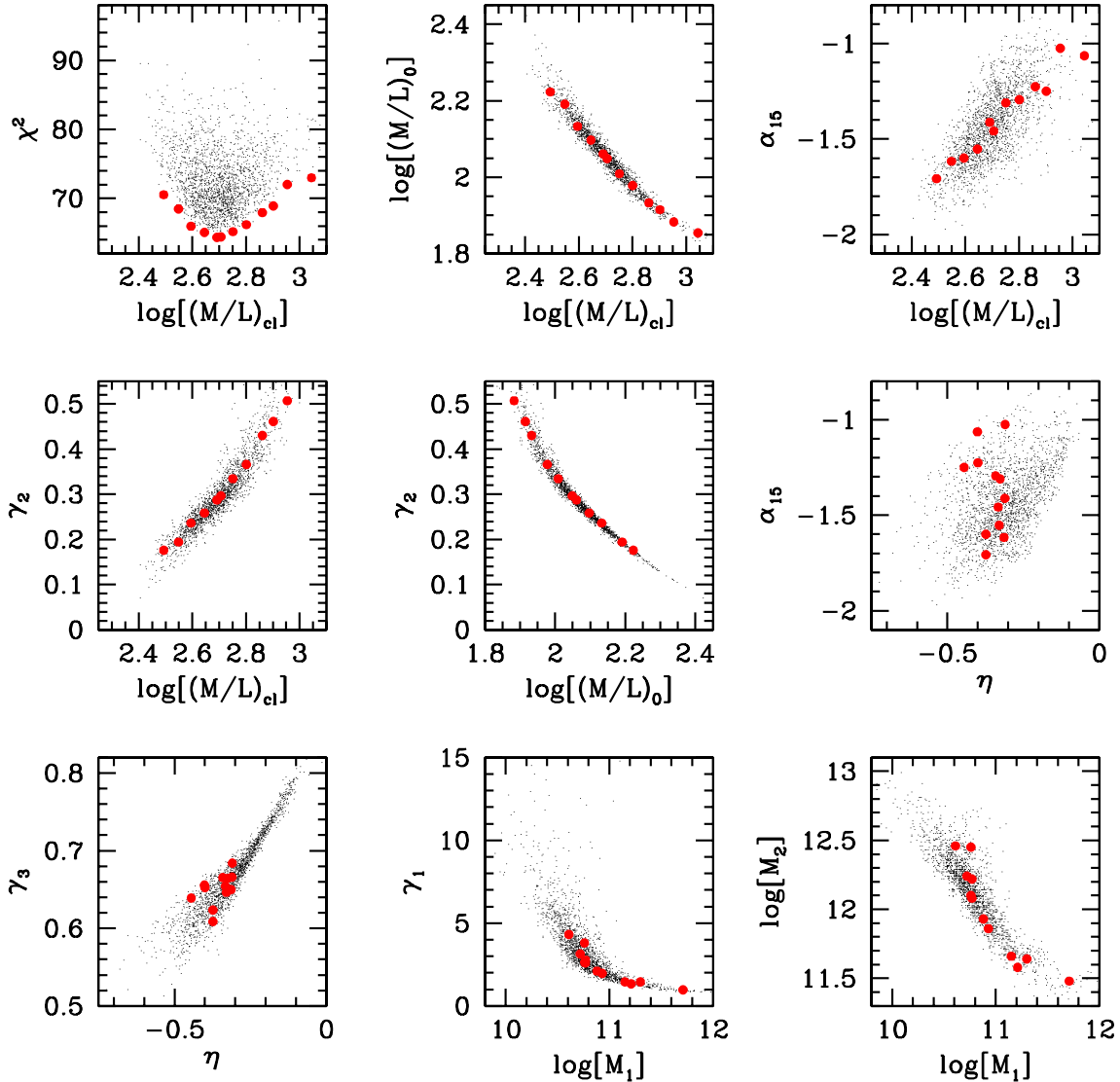
Figure 2. Correlations between various CLF parameters of the 2000 samples in the $\sigma_8 = 0.9$ MCMC (thin dots). The thick, solid dots correspond to the 12 CLF models listed in Table 1 which represent the best-fitting models in 12 intervals of cluster mass-to-light ratio, $(M/L)_{cl}$. Masses and mass-to-light ratios are in units of $h^{-1} M_\odot$ and h (M/L)$_\odot$, respectively.

redshift surveys (hereafter MGRSs). These have the following advantages:

   (i) we know exactly the *true* abundances and the *true* radial distributions;

   (ii) we can model the effects of various observational biases; and

   (iii) we can use exactly the same host/satellite selection criteria as for the 2dFGRS, making the comparison with the data straightforward.

To construct MGRSs two ingredients are required: a distribution of dark matter haloes and a description of how galaxies of different luminosity occupy haloes of different mass. For the former we use large numerical simulations, and for the latter the CLF. Ideally we would construct a MGRS for each of the 2000 models in our MCMC. As this is computationally too expensive we adopt an alternative method. We determine the minimum and maximum values of $(M/L)_{cl}$ in the entire MCMC, and split the interval into 12 equally sized logarithmic bins. For each of these bins we then determine the model that yields the lowest χ^2. The parameters of the resulting

12 models are listed in Table 1, and are indicated by thick solid dots in Fig. 2. For each of these 12 CLFs we construct a MGRS, which we use to assess the uncertainties on the abundances and radial distribution of satellite galaxies owing to the uncertainties in $(M/L)_{cl}$. In what follows we refer to the MGRS based on CLF n as Mn, where n is the ID listed in column (1) of Table 1.

We choose $(M/L)_{cl}$ as our main parameter to sample parameter space because (i) this parameter has direct observational constraints (e.g. Carlberg et al. 1996; Fukugita, Hogan & Peebles 1998; Bahcall et al. 2000), (ii) previous studies based on the CLF have shown that varying $(M/L)_{cl}$ has important implications for the pairwise peculiar velocities of galaxies (Yang et al. 2004a) and the abundance of large groups (Yang et al. 2004b) and (iii) it is strongly correlated with several of the other CLF parameters (see Fig. 2).

3.1 Numerical simulations

The distribution of dark matter haloes is obtained from a set of large *N*-body simulations (dark matter only) for a ΛCDM 'concordance'
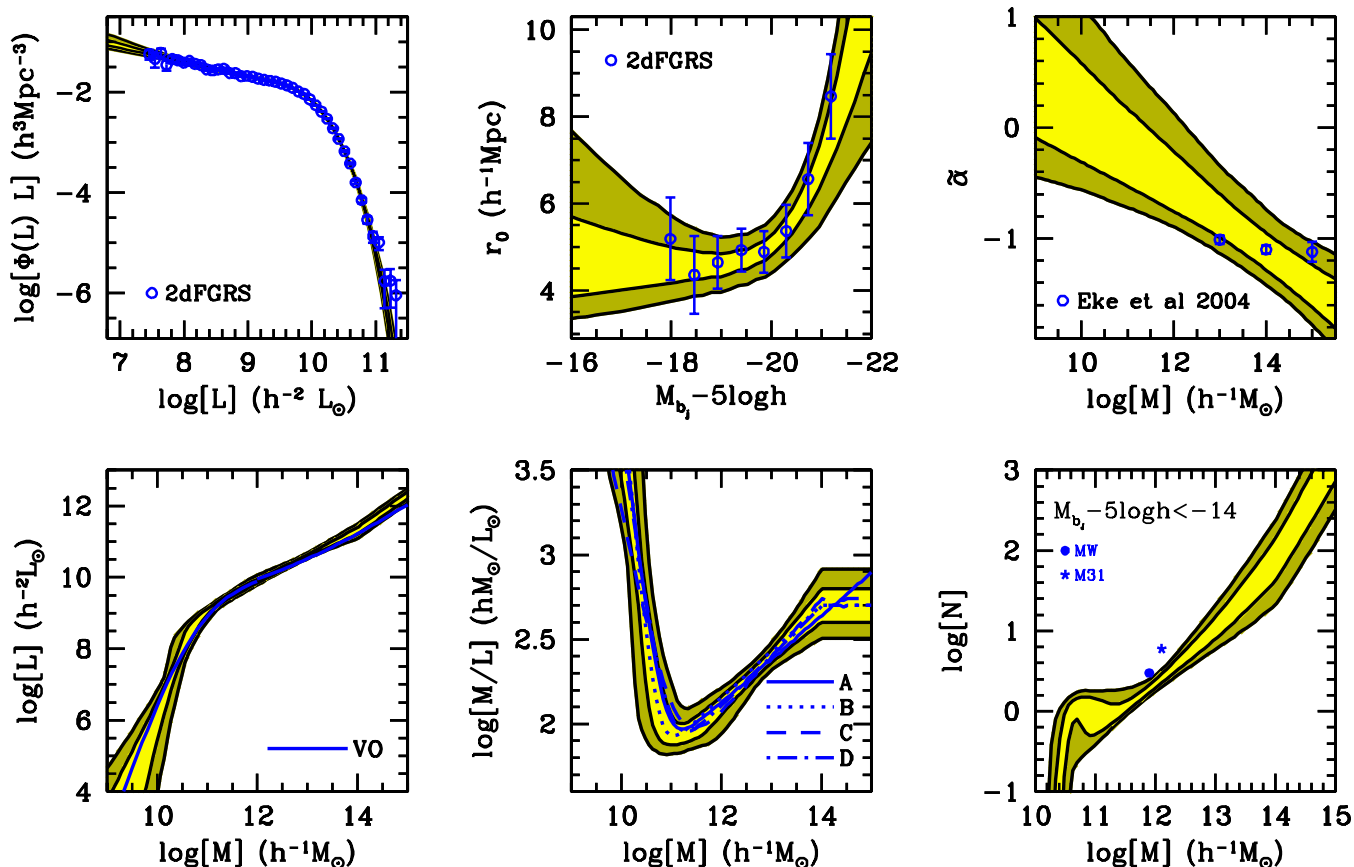
Figure 3. Posterior constraints on a number of quantities computed from the $\sigma_8 = 0.9$ MCMC. The contours show the 68 and 99 per cent confidence limits from the marginalized distribution. Upper left-hand panel: the galaxy luminosity function; open circles with error bars correspond to the 2dFGRS data from Madgwick et al. (2002). Upper middle panel: galaxy–galaxy correlation lengths as function of absolute magnitude; open circles with error bars correspond to the 2dFGRS data from Norberg et al. (2002a). Upper right-hand panel: the faint-end slope of the CLF as a function of halo mass; open circles with error bars correspond to the 2dFGRS data from Eke et al. (2004). Lower left-hand panel: the total luminosity per halo as a function of halo mass. The solid grey line (blue in the electronic version) corresponds to the model of Vale & Ostriker (2004), and is shown for comparison. Lower middle panel: the average mass-to-light ratio as a function of halo mass. The various lines (blue in the electronic version) correspond to models A–D of van den Bosch et al. (2003a), as indicated. Lower right-hand panel: the average number of galaxies with $M_{b_J} - 5 \log h < -14$ per halo as a function of halo mass. See text for a detailed discussion.

cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$ and $\sigma_8 = 0.9$. In this paper we use two simulations with $N = 512^3$ particles each, which are described in more detail in Yang et al. (2004a) and Jing & Suto (2002). The simulations have periodic boundary conditions and box sizes of $L_{box} = 100\,h^{-1}$ Mpc (hereafter L_{100}) and $L_{box} = 300\,h^{-1}$ Mpc (hereafter L_{300}). We follow Yang et al. (2004a) and replicate the L_{300} box on a $4 \times 4 \times 4$ grid. The central $2 \times 2 \times 2$ boxes, are replaced by a stack of $6 \times 6 \times 6$ L_{100} boxes, and the virtual observer is placed in the centre (see fig. 11 in Yang et al. (2004a)). This stacking geometry circumvents incompleteness problems in the mock survey owing to insufficient mass resolution of the L_{300} simulations, and allows us to reach the desired depth of $z_{max} = 0.15$ in all directions.

Dark matter haloes are identified using the standard friends-of-friends algorithm (Davis et al. 1985) with a linking length of 0.2 times the mean interparticle separation. Unbound haloes, or haloes with less than 10 particles, are removed from the sample. In Yang et al. (2004a) we have shown that the resulting halo mass functions are in excellent agreement with the analytical halo mass function given by Sheth et al. (2001) and Sheth & Tormen (2002).

3.2 Halo occupation numbers and luminosities

Owing to the mass resolution of the simulations and the completeness limit of the 2dFGRS we adopt a minimum galaxy luminosity of $L_{min} = 10^7\,h^{-2}\,L_\odot$ throughout. The *mean* number of galaxies with $L \geqslant L_{min}$ that resides in a halo of mass M is given by equation (13). In order to Monte Carlo sample occupation numbers for individual haloes one requires the full probability distribution $P(N|M)$ (with N an integer) of which $\langle N \rangle_M$ gives the mean, i.e.

$$\langle N \rangle_M = \sum_{N=0}^{\infty} N\,P(N|M). \tag{14}$$

We use the results of Kravtsov et al. (2004a), who has shown that the number of *subhaloes* follows a Poisson distribution. In what follows we differentiate between satellite galaxies, which we associate with these dark matter subhaloes, and central galaxies, which we associate with the host halo (cf. Vale & Ostriker 2004). The total number of galaxies per halo is the sum of N_{cen}, the number of central galaxies which is either one or zero, and N_{sat}, the (unlimited) number of satellite galaxies. We assume that N_{sat} follows a Poisson distribution and require that $N_{sat} = 0$ whenever $N_{cen} = 0$. The halo

occupation distribution is thus specified as follows: if $\langle N \rangle_M \leqslant 1$ then $N_{sat} = 0$ and N_{cen} is either zero (with probability $P = 1 - \langle N \rangle_M$) or one (with probability $P = \langle N \rangle_M$). If $\langle N \rangle_M > 1$ then $N_{cen} = 1$ and N_{sat} follows the Poisson distribution

$$P(N_{sat}|M) = e^{-\mu} \frac{\mu^{N_{sat}}}{N_{sat}!}, \tag{15}$$

with $\mu = \langle N_{sat} \rangle_M = \langle N \rangle_M - 1$.

We follow Yang et al. (2004a) and van den Bosch et al. (2004b) and assume that the central galaxy is the brightest galaxy in each halo. Its luminosity is drawn from $\Phi(L|M)$ with the restriction that $L > L_1$ with L_1 defined by

$$\int_{L_1}^{\infty} \Phi(L|M)\, dL = 1. \tag{16}$$

The luminosities of the satellite galaxies are also drawn at random from $\Phi(L|M)$, but with the restriction $L_{min} < L < L_1$.

3.3 Assigning galaxies their phase-space coordinates

Next we assign all galaxies a position and velocity within their halo. Ideally we would associate the central galaxy with the parent halo, and satellite galaxies with the dark matter subhaloes. However, as our numerical simulations do not resolve these subhaloes, we have to assign the positions and velocities of (satellite) galaxies 'manually'. We assume that each dark matter halo has an NFW density distribution with virial radius r_{vir}, characteristic scale radius r_s, and concentration $c = r_{vir}/r_s$. Throughout this paper we compute halo concentrations as a function of halo mass using the relation given by Eke, Navarro & Steinmetz (2001), properly accounting for our definition of halo mass. The 'central' (brightest) galaxy in each halo is assumed to be located at the halo centre, which we associate with the position of the most bound particle. Satellite galaxies are assumed to follow a radial number density distribution given by

$$n_{sat}(r) \propto \left(\frac{r}{\mathcal{R} r_s} \right)^{-\alpha} \left(1 + \frac{r}{\mathcal{R} r_s} \right)^{\alpha - 3} \tag{17}$$

(limited to $r \leqslant r_{vir}$) with α and $\mathcal{R}$ being two free parameters. Unless specifically stated otherwise we adopt $\alpha = \mathcal{R} = 1$ for which the number density distribution of satellite galaxies exactly follows the dark matter mass distribution.

Finally, peculiar velocities are assigned as follows. We assume that the 'central' galaxy is located at rest with respect to its halo, and set its peculiar velocity equal to the mean halo velocity. Satellite galaxies are assumed to be in a steady-state equilibrium within the dark matter potential well with an isotropic distribution of velocities with respect to the halo centre (see van den Bosch et al. 2004b for details). As shown by Diemand et al. (2004) this is a good approximation for dark matter subhaloes.

3.4 Creating mock surveys

The 2dFGRS uses a multifibre spectrograph to obtain redshifts. However, because of the physical size of the fibres, when two galaxies are closer than ~30 arcsec in projection only one of them can be targeted. Furthermore, owing to clustering, some areas on the sky contain more galaxies within a single two-degree field than the available number of fibres. By using a sophisticated tiling strategy these problems are largely overcome, yielding a fairly uniform sampling rate. Nevertheless, some spatial non-uniformities remain. In addition, fainter galaxies yield noisier spectra, and therefore less accurate redshifts. All these effects combined result in a complete-

ness which depends on both the position in the sky and on the apparent magnitude. The 2dFGRS team has constructed maps that parametrize this position and magnitude-dependent completeness (Colless et al. 2001; Norberg et al. 2002b), and which facilitate a correction for these effects in our MGRSs. However, as it turns out, the completeness depends also on the angular separation, θ, between galaxy *pairs*. This is largely owing to the problem of fibre collisions, which has not been completely corrected for by the tiling strategy. Finally, Norberg et al. (2002b) have shown that the *parent* catalogue of the 2dFGRS, the APM catalogue (Maddox et al. 1990), is only 91 per cent complete. Using this information we mimic the various observational selection and completeness effects in the 2dFGRS using the following steps.

(i) We define a (α, δ)-coordinate frame with respect to the virtual observer at the centre of the stack of simulation boxes, and remove all galaxies that are not located in the areas equivalent to the NGP and SGP regions of the 2dFGRS.

(ii) For each galaxy we compute the redshift as 'seen' by the virtual observer. We take the observational velocity uncertainties into account by adding a random velocity drawn from a Gaussian distribution with dispersion $85\ km\ s^{-1}$ (Colless et al. 2001).

(iii) For each galaxy we compute the apparent magnitude according to its luminosity and distance, to which we add a rms error of 0.15 mag (Colless et al. 2001; Norberg et al. 2002b). As galaxies in the 2dFGRS were pruned by apparent magnitude *before* a K correction was applied, we proceed as follows: we first apply a negative K correction, then select galaxies according to the position-dependent magnitude limit (obtained using the apparent magnitude limit masks provided by the 2dFGRS team), and finally K correct the magnitudes back to their rest-frame b_J band. Throughout we use the type-dependent K corrections given in Madgwick et al. (2002).

(iv) For each galaxy we compute the redshift as 'seen' by the virtual observer. We take the observational velocity uncertainties into account by adding a random velocity drawn from a Gaussian distribution with dispersion $85\ km\ s^{-1}$ (Colless et al. 2001), and remove those galaxies not in the redshift range $0.01 < z < 0.15$.

(v) To take account of the completeness level of the 2dFGRS parent catalogue (Norberg et al. 2002b) we randomly remove 9 per cent of all galaxies.

(vi) To take account of the position- and magnitude-dependent completeness of the 2dFGRS, we randomly sample each galaxy using the completeness masks provided by the 2dFGRS team.

The construction of the MGRS described above is identical to the method used by Yang et al. (2004a,b), Wang et al. (2004), and van den Bosch et al. (2004b). As discussed above, the level of completeness of the 2dFGRS depends on position, apparent magnitude and pair-separation θ. In constructing our MGRSs we corrected for the formed two effects, using detailed completeness maps provided by the 2dFGRS team, but no correction was made for the pair-separation dependency. However, for the purpose of investigating the radial distribution of satellite galaxies this θ-dependent incompleteness cannot be ignored.

To quantify the pair-separation incompleteness Hawkins et al. (2003) computed the ratio $r(\theta) \equiv [1 + w_p(\theta)]/[1 + w_z(\theta)]$ between the angular correlation functions of the 2dFGRS parent catalogue, $w_p(\theta)$, and that of the final redshift survey, $w_z(\theta)$. $r(\theta)$ is therefore a direct measure of the fraction of pairs with separation θ in the parent catalogue divided by the same fraction but in the final redshift survey. Whenever $r(\theta)$ differs from unity it indicates that the incompleteness of the redshift survey (with respect to the parent catalogue) depends on pair separation, where values larger than
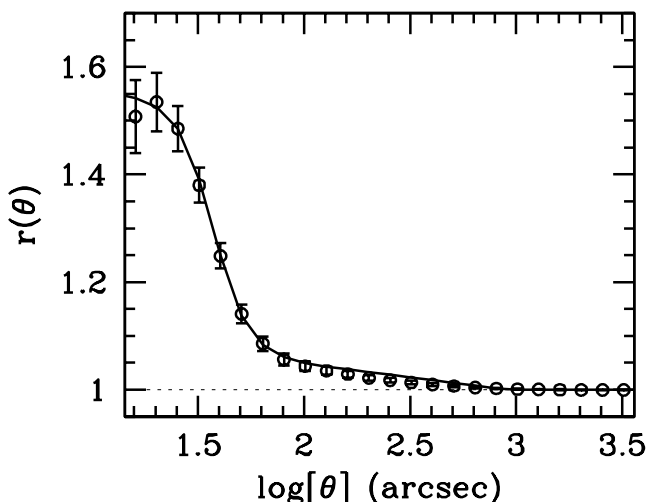
Figure 4. The solid line corresponds to the $r(\theta)$ (see text for definition) obtained from the 2dFGRS by Hawkins et al. (2003). Note that r is significantly larger than unity for $\theta \lesssim 100$ arcsec, indicating that the 2dFGRS has missed a significant fraction of close pairs (largely owing to fibre collisions). The open circles with error bars correspond to the $r(\theta)$ obtained for MGRS M6, after we have removed close pairs. The good agreement with the results from the 2dFGRS indicates that we have constructed a MGRS with the same deficiency of close pairs as in the data. Note that this does not correct for a possible close-pair incompleteness in the *parent* catalogue of the 2dFGRS (see Section 5.2).

unity indicate a pair *deficiency*. Hawkins et al. found r to be significantly larger than unity for $\theta \lesssim 60$ arcsec. In an attempt to model this 'close-pair deficiency' in our MGRSs we proceed as follows. We follow the same procedure as described above, but we skip step (iv); i.e. we do not correct the MGRS for the incompleteness of the parent catalogue. After having corrected for the position- and magnitude-dependent incompleteness, we compute the angular separations θ between all galaxy pairs and remove galaxies based on a probability $p(\theta)$, which we tune (by trial and error) so that we reproduce the $r(\theta)$ obtained by Hawkins et al. (2003). As a last step we then remove a number of galaxies completely at random such that, together with the galaxies removed because of their angular separation to neighbours, we have removed 9 per cent of all mock galaxies. This mimics the incompleteness level of the parent catalogue.[2] The solid line in Fig. 4 shows $r(\theta)$ for the 2dFGRS as obtained by Hawkins et al. (2003), while open circles with error bars correspond to the $r(\theta)$ obtained from a comparison of the angular correlation functions in MGRS M6 before and after the removal of close pairs. The good agreement with the 2dFGRS results demonstrates that we have managed to construct MGRSs with the same deficiency of close pairs as present in the real 2dFGRS. All MGRSs discussed in this paper have been similarly corrected for this close-pair incompleteness.

4 SELECTING HOST AND SATELLITE GALAXIES

A galaxy is considered a potential host galaxy if it is at least f_h times brighter than any other galaxy within a volume specified by $R_p < R_h$

----

[2] Note that the sequential order in which the various completeness corrections have been applied is not entirely correct. However, numerous tests have shown that this has an absolutely negligible effect on our results.

and $|\Delta V| < (\Delta V)_h$. Here R_p is the separation projected on the sky at the distance of the candidate host, and ΔV is the line-of-sight velocity difference. Around each potential host galaxy, satellite galaxies are defined as those galaxies that are at least f_s times fainter than their host and located within a volume with $R_p < R_s$ and $|\Delta V| < (\Delta V)_s$ (here ΔV is the line-of-sight velocity difference between the potential satellite galaxy and the host galaxy under consideration). Host galaxies with zero satellite galaxies are removed from the list of hosts.

In total, the selection of hosts and satellites thus depends on six free parameters: R_h, $(\Delta V)_h$ and f_h to specify the population of host galaxies, and R_s, $(\Delta V)_s$ and f_s to specify the satellite galaxies. These parameters also determine the number of interlopers (defined as a galaxy not physically associated with the halo of the host galaxy) and non-central hosts (defined as a host galaxy that is not the brightest, central galaxy in its own halo). Minimizing the number of interlopers requires sufficiently small R_s and $(\Delta V)_s$. Minimizing the number of non-central hosts requires one to choose R_h, $(\Delta V)_h$ and f_h sufficiently large. Of course, each of these restrictions dramatically reduces the number of both hosts and satellites, making the statistical estimates more and more noisy.

In van den Bosch et al. (2004b) we used our MGRS to optimize these selection criteria, aiming for large numbers of hosts and satellites, a small fraction of interlopers, and a small fraction of non-central hosts. It was shown that an adaptive selection criterion, for which R_h, $(\Delta V)_h$ and R_s are made dependent of the luminosity of the host candidate, was most successful. Motivated by these findings we adopt: $f_h = f_s = 1$, $(\Delta V)_h = 1000 \, \sigma_{200} \, \mathrm{km \, s^{-1}}$, $(\Delta V)_s = 2000 \, \mathrm{km \, s^{-1}}$, $R_h = 0.8 \, \sigma_{200} \, h^{-1} \, \mathrm{Mpc}$ and $R_s = 0.15 \, \sigma_{200} \, h^{-1} \, \mathrm{Mpc}$. Here $\sigma_{200} = \sigma_{sat}(L_{host})/(200 \, \mathrm{km \, s^{-1}})$ is the velocity dispersion of satellite galaxies around a host halo of luminosity L_{host} in units of $200 \, \mathrm{km \, s^{-1}}$. Following van den Bosch et al. (2004b) we adopt

$$\log[\sigma_{sat}(L_{host})] = \log\sigma_{10} + a_1 \log L_{10} + a_2 (\log L_{10})^2. \quad (18)$$

Here $L_{10} = L_{host}/10^{10} \, h^{-2} \, \mathrm{L_\odot}$ and $\sigma_{10} = \sigma_{sat}(L_{10})$. We set $\sigma_{10} = 200 \, \mathrm{km \, s^{-1}}$, $a_1 = 0.5$ and $a_2 = 0.1$, which is close to the parameters obtained for the 2dFGRS by van den Bosch et al. (2004b). Note that details regarding the selection criteria of host and satellite galaxies are not very important as long as the data and model (MGRS) are treated in the same way.

For the data we use the final, public data release from the 2dFGRS, restricting ourselves only to galaxies with redshifts $0.01 \leqslant z \leqslant 0.15$ in the North Galactic Pole and South Galactic Pole subsamples with a redshift quality parameter $q \geqslant 3$. This leaves a grand total of 146 735 galaxies with a typical rms redshift error of $85 \, \mathrm{km \, s^{-1}}$ (Colless et al. 2001). Absolute magnitudes for galaxies in the 2dFGRS are computed using the K corrections of Madgwick et al. (2002). Applying our host–satellite selection criterion yields 8737 host galaxies and 13 738 satellite galaxies.

5 THE RADIAL DISTRIBUTION OF SATELLITE GALAXIES

The hatched histograms in Fig. 5 show the *projected*, radial, number density distribution, $\bar{\Sigma}_{sat}(R_p)$ (in h^2/Mpc^2/host) of 2dFGRS satellite galaxies around hosts in three different host magnitude bins. Note that the $\bar{\Sigma}_{sat}(R_p)$ reveal both an inner and an outer break. The outer break radius increases with increasing host luminosity, and is mainly owing to our selection criteria. The inner break radius, however, indicates a pronounced absence of satellite galaxies at small projected radii from the host galaxy. To interpret these findings in terms of the actual three-dimensional number density distribution of
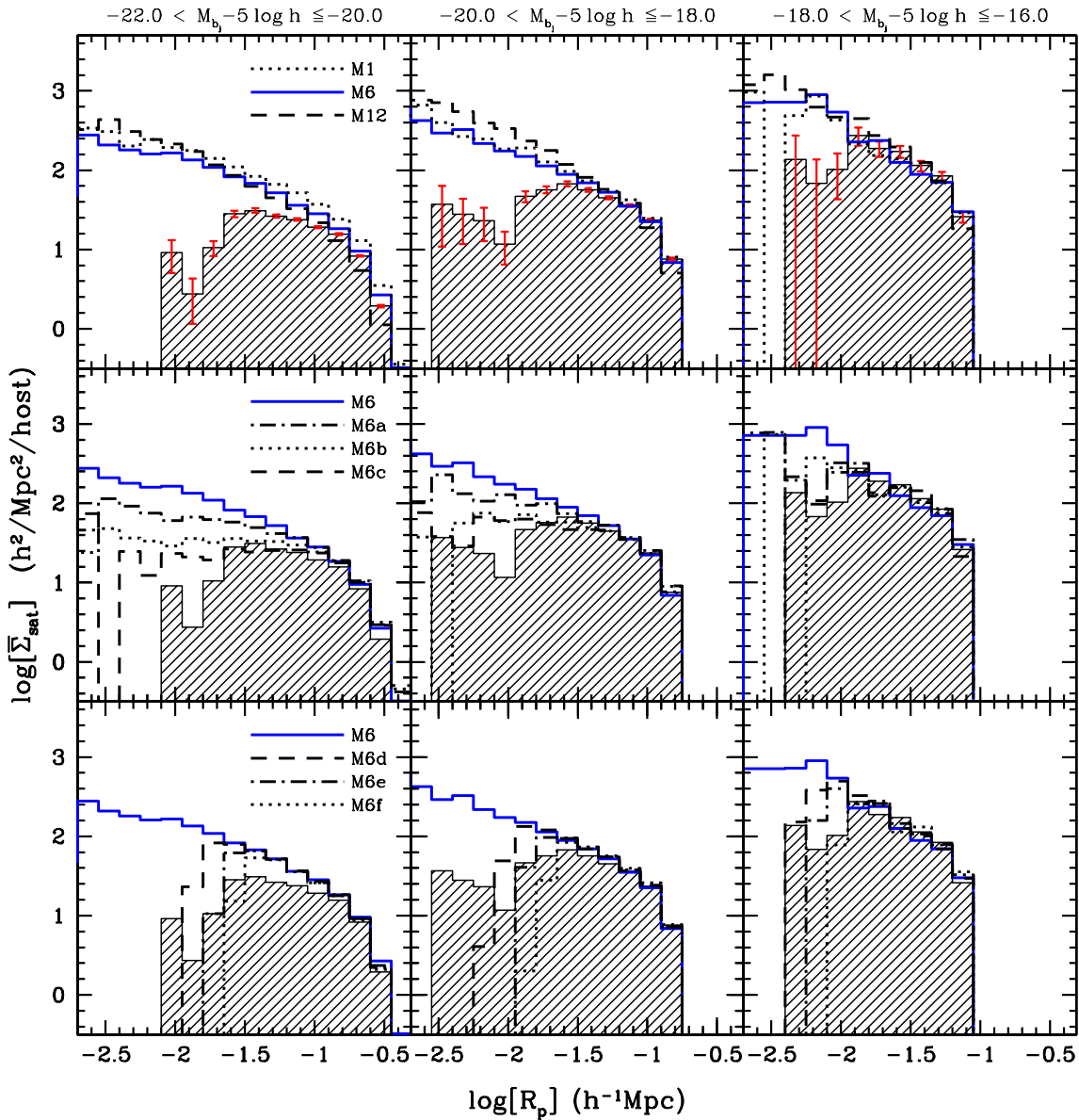
Figure 5. The projected number density distributions of satellite galaxies as a function of projected radius. Results are shown for three different bins in host–galaxy luminosity, indicated at the top of each column. In each panel, the hatched histogram corresponds to the 2dFGRS data, while the various lines have been obtained from the MGRSs. The error bars (only shown in the upper panels for clarity) indicate Poisson errors in the numbers of host and satellite galaxies. The upper panels compare the 2dFGRS data to three MGRSs (M1, M6 and M12) that span the entire range in cluster mass-to-light ratio (see Table 1 for parameters). The middle row of panels compare the 2dFGRS data to MGRSs based on CLF 6, but in which the intrinsic, spatial distribution of satellite galaxies is modelled with a constant number density core (equation 17 with $\alpha = 0$) with $R = 1$ (M6a), 2 (M6b) and 3 (M6c). For comparison the fiducial MGRS M6, for which $\alpha = R = 1$ is also shown (solid lines). Finally, the lower panels compare the 2dFGRS data to MGRSs based on CLF 6, but in which we have corrected for the size-projection incompleteness using $R_{10} = 10\ h^{-1}$ kpc (M6d), $R_{10} = 15\ h^{-1}$ kpc (M6e) and $R_{10} = 20\ h^{-1}$ kpc (M6f). See text for detailed discussion.

satellite galaxies we compare these findings to those obtained from our MGRSs.

The three non-hatched histograms in the upper panels of Fig. 5 depict the $\bar{\Sigma}_{\rm sat}(R_{\rm p})$ as obtained from MGRSs M1 (dotted lines), M6 (solid lines) and M12 (dashed lines). Note that all three MGRSs yield virtually identical radial number density distributions. In fact, this holds for all 12 MGRSs listed in Table 1: for clarity, however, only three are shown. Although the projected number density distributions of the MGRSs nicely match the 2dFGRS data at large projected radii, they severely overpredict $\bar{\Sigma}_{\rm sat}(R_{\rm p})$ at small $R_{\rm p}$. Rather than a pronounced inner break radius, $\bar{\Sigma}_{\rm sat}(R_{\rm p})$ increases continu-

ously down to the smallest radii shown. Especially for the brightest host galaxies the discrepancy between MGRSs and 2dFGRS is dramatic. Either satellite galaxies are spatially antibiased with respect to the dark matter, or the 2dFGRS has somehow missed a large number of satellite galaxies at small projected radii from their host galaxies. We investigate both options below.

Another characteristic of the projected number density distributions shown is that $\bar{\Sigma}_{\rm sat}$ is higher around fainter host galaxies. The surface density of satellite galaxies scales as $\bar{\Sigma}_{\rm sat} \propto \langle N \rangle_{\rm s}/R_{\rm vir}^2$, with $\langle N \rangle_{\rm s}$ the average number of satellites in a halo associated with the host galaxy, and $R_{\rm vir}$ the corresponding virial radius. To a good

approximation, $\langle N \rangle_s \propto M^\alpha$, with M the halo mass and $\alpha \simeq 0.9$ (e.g. Kravtsov et al. 2004a). Writing the relation between host luminosity and halo mass as a simple power law, $L_{host} \propto M^\beta$, and using that $R_{vir} \propto M_{vir}^{1/3}$, one thus predicts that $\bar{\Sigma}_{sat} \propto L_{host}^{0.23/\beta}$. Given that $\beta > 0$ one thus expects that the projected number density decreases with decreasing L_{host}, opposite to what is seen in Fig. 5. The reason for this apparent disagreement is that we only consider host galaxies with at least one satellite. This means that, because of our selection criteria, $\langle N \rangle_s$ deviates from the pure power law at the low-mass end, where it assymptotes to $\langle N \rangle_s = 1$ (i.e. $\alpha = 0.0$). In this case $\bar{\Sigma}_{sat} \propto L_{host}^{-0.67/\beta}$ and $\bar{\Sigma}_{sat}$ increases with decreasing L_{host}, as observed.[3]

5.1 Spatial antibias

When constructing the MGRSs we have thus far assumed that satellite galaxies follow the same number density distribution as dark matter particles (i.e. equation 17 with $\alpha = \mathcal{R} = 1.0$). The resulting overestimate of the projected number density of satellite galaxies at small R_p, however, might indicate that $\alpha < 1.0$ and/or that $\mathcal{R} > 1$. To test these ideas we construct three MGRSs using CLF 6. These MGRSs, termed M6a, M6b and M6c, all have $\alpha = 0.0$, and only differ in the value of $\mathcal{R}$, which is set to 1.0, 2.0 and 3.0, respectively. The projected number density distributions of their satellite galaxies are shown in the middle row of panels of Fig. 5. For comparison we also plot the $\bar{\Sigma}_{sat}(R_p)$ of M6 (solid lines). As expected, lowering α and increasing $\mathcal{R}$ both result in a decrease of $\bar{\Sigma}_{sat}$ at small R_p. In fact, MGRS M6c fits most of the 2dFGRS data reasonably well, although it fails to reproduce the pronounced inner break evident in the data.

Taking these results at face value seems to indicate that satellite galaxies are spatially antibiased with respect to the dark matter mass distribution on small scales (20–30 per cent of the virial radius). This is in wonderful agreement with high-resolution numerical N-body simulations, which reveal a very similar spatial antibias for dark matter subhaloes (Ghigna et al. 1998; Colin et al. 1999; Okamoto & Habe 1999; Ghigna et al. 2000; Springel et al. 2001; De Lucia et al. 2004; Diemand et al. 2004), but inconsistent with observations of the number density distribution of galaxies in clusters (Beers & Tonry 1986; Carlberg et al. 1997; van der Marel et al. 2000; Diemand et al. 2004; Lin et al. 2004) and with predictions of semi-analytical models of galaxy formation (Springel et al. 2001; Diaferio et al. 2001; Gao et al. 2004a).

5.2 Observational projection effects

This puzzling inconsistency calls for a closer examination of the selection effects in the 2dFGRS. Could the absence of satellite galaxies at small projected radii from their host galaxies be owing to selection effects? Note that it cannot be explained as owing to fibre collisions. As shown in Fig. 4, and discussed in Section 3, these effects are accurately accounted for in our MGRSs. One effect that has not yet been accounted for, however, is related to the fact that galaxies have a finite size, which causes them to overlap in projection. Whenever this happens only the brightest of the two will be recognized as a

galaxy and included in the survey. Indeed, as shown by Cole et al. (2001) using a comparison of the 2dFGRS with the Two Micron All Sky Survey (2MASS), about 4.5 per cent of all galaxies is missed in the 2dFGRS parent catalogue because of merged or close images (explaining about half of the incompleteness; the other half being because of incorrect star–galaxy separation).

In order to take this size-projection effect into account we proceed as follows. We model the characteristic size of a galaxy as

$$R_{gal} = R_{10} \left(\frac{L}{10^{10} \, h^{-2} \, L_\odot} \right)^\zeta \tag{19}$$

and define the critical projection angle $\theta_{max} = R_{gal}/D_A$, with D_A the angular distance of the galaxy. Based on the data on disc galaxies in Courteau et al. (2003) we adopt $\zeta = 1/3$, but we note that this assumption is not very important. For example, we have also experimented with $\zeta = 1/2$ finding very similar results as those described here.

We construct three MGRSs in which we take the size-projection effect into account using $R_{10} = 10 \, h^{-1}$ kpc (M6d), $R_{10} = 15 \, h^{-1}$ kpc (M6e), and $R_{10} = 20 \, h^{-1}$ kpc (M6f). We follow the same procedure as described in Section 3, except that this time, after having corrected for the position- and magnitude-dependent incompleteness, we compute the angular separations θ between all galaxy pairs and remove the faintest galaxy from all pairs for which $\theta < \theta_{max}$. Typically this removes about 2–4 per cent of all galaxies, which is comparable to the incompleteness level in the 2dFGRS parent catalogue owing to this effect as found by Cole et al. (2001). Next we correct for the fibre collisions, using the same trial-and-error method as described in Section 3. Finally we remove a number of galaxies completely at random to bring the total fraction of removed galaxies to 9 per cent.

The lower panels Fig. 5 plot the $\bar{\Sigma}_{sat}(R_p)$ of MGRSs M6d (dashed lines), M6e (dot-dashed lines) and M6f (dotted lines). Note that these MGRSs have $\alpha = \mathcal{R} = 1.0$, as for our fiducial MGRSs. For comparison, the solid lines plot the projected number density distributions of MGRS M6, for which no correction for the size-projection effect has been applied. Clearly, the size-projection effect as modelled here has a dramatic impact on $\bar{\Sigma}_{sat}(R_p)$ at small projected radii. In fact, it reproduces the inner cut-off observed for the 2dFGRS data, indicating that, as first noticed by Cole et al. (2001), the 2dFGRS lacks a significant fraction of close galaxy pairs owing to projection effects. The dramatic impact of this size-projection effect on $\bar{\Sigma}_{sat}(R_p)$ owes to the fact that in our MGRSs the central galaxy in each halo is always the brightest galaxy in that halo. Most of the pairs with small separations (i.e. those that would overlap in projection) are between such a central galaxy and one of the inner satellites. As the latter is always the fainter of the two, the number density of satellites at small projected separations from the central galaxy is drastically reduced.

Unfortunately, unless we can model this particular incompleteness in more detail as done here (i.e. have independent constraints on R_{10} and ζ), we cannot use the 2dFGRS data to meaningfully constrain the actual radial density distribution of satellite galaxies. Our models suggest that the data is consistent with no spatial (anti)bias as long as the size-projection effects are taken into account, but we cannot rule out that $\alpha = 0.0$ and/or that $\mathcal{R} > 1.0$. For example, models with both $\alpha = 0.0$ and in which we correct for the size-projection effect are virtually indistinguishable from the same models but with $\alpha = 1.0$. It remains to be seen whether the higher spatial resolution of the Sloan Digital Sky Survey (SDSS) will allow a more in-depth investigation of the radial distribution of satellite galaxies.

---

[3] Although this argument is based on the assumption of a volume-limited sample, using our MGRSs we have verified that the host/satellite systems selected from a flux-limited sample reveal the same behavior, with $\langle N \rangle_s$ changing smoothly from a power law with slope ~ 0.8 at the bright end to a constant at the faint end.

6 THE ABUNDANCE OF HOST AND SATELLITE GALAXIES

We now focus on the abundances of host and satellite galaxies. For the 2dFGRS we obtain host and satellite fractions of $N_{host}/N_{tot} = 0.060$ (0.060, 0.059) and $N_{sat}/N_{tot} = 0.094$ (0.096, 0.092), where the numbers in parentheses indicate the fractions obtained using only the NGP and SGP data, respectively. We compare these numbers with those obtained from the MGRSs. The solid triangles in the upper panels of Fig. 6 plot N_{host}/N_{tot}, N_{sat}/N_{tot} and N_{sat}/N_{host} as a function of $(M/L)_{cl}$ for the 12 MGRSs listed in Table 1. The grey, horizontal bars indicate the 2dFGRS results. Clearly, the fraction

of satellite galaxies in the MGRS is much too high compared to the 2dFGRS, independent of $(M/L)_{cl}$. The fraction of host galaxies is in reasonable agreement with the 2dFGRS (though somewhat too high), but only for $(M/L)_{cl} \lesssim 650\,h$ $(M/L)_{\odot}$. Matching the number of satellites per host, however, favours much higher values for $(M/L)_{cl}$. In short, none of the 12 MGRS can simultaneously match the abundances of host and satellite galaxies in the 2dFGRS.

The open circles in the upper panels of Fig. 6 correspond to the same 12 MGRSs, but this time corrected for the size-projection effects discussed in Section 5.2 using $R_{10} = 0.015\,h^{-1}$ Mpc and $\zeta = 1/3$. This lowers the overall abundances of both hosts and (mainly) satellites, but still, none of the 12 MGRS can simultaneously match
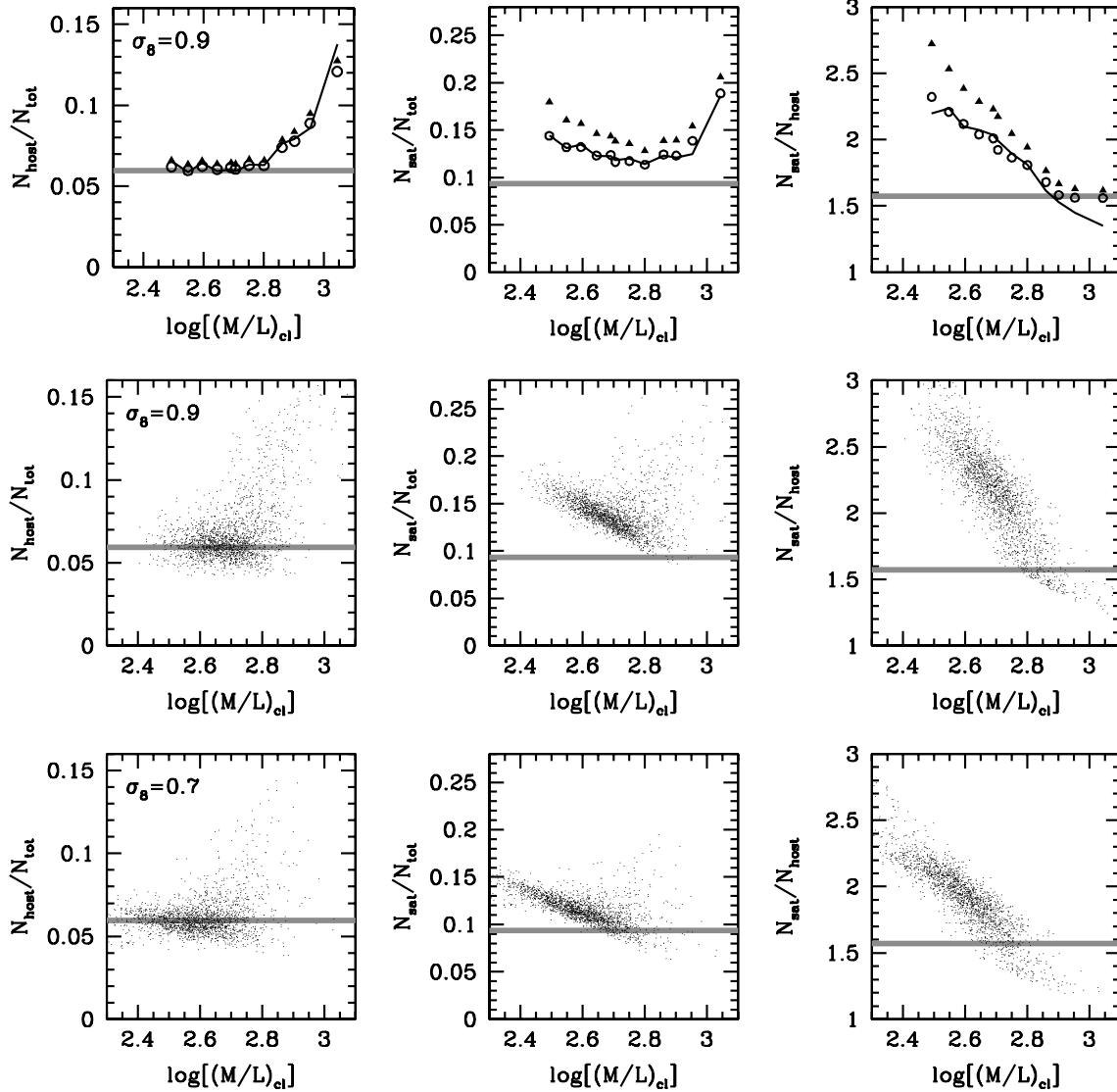
Figure 6. Abundances of host and satellite galaxies. From left to right, the different panels in a given row plot the ratios N_{host}/N_{tot}, N_{sat}/N_{tot} and N_{sat}/N_{host}, all as a function of the cluster mass-to-light ratio $(M/L)_{cl}$ (in units of h $(M/L)_{\odot}$). Grey, horizontal bars indicate the ratios obtained from the 2dFGRS. Upper panels: solid triangles and open circles indicate the ratios obtained from the MGRSs for the 12 CLFs listed in Table 1 without and with correction for the size-projection effects, respectively. Error bars, computed assuming Poisson errors on N_{tot}, N_{host} and N_{sat}, are smaller than the symbols, and are therefore not shown. The solid line indicates the ratios obtained from the analytical estimates (equations 20–25) after taking account of the correction factors f_{host} and f_{sat} (see text). The good agreement with the open circles indicates that these analytical estimates can be used to compute the abundances of host and satellite galaxies for any CLF model without the need to construct a detailed MGRS. Middle panels: the ratios N_{host}/N_{tot}, N_{sat}/N_{tot} and N_{sat}/N_{host} for all 2000 samples in the $\sigma_8 = 0.9$ MCMC, obtained using the approximate, analytical method. Note that none of these models can simultaneously match all three ratios obtained from the 2dFGRS. Lower panels: same as middle panels, but this time showing the results for the $\sigma_8 = 0.7$ MCMC. Note that at around $\log[(M/L)_{cl}] \simeq 2.75$ these model predictions simultaneously match all three 2dFGRS ratios.

the abundances of host and satellite galaxies in the 2dFGRS. Although the host fraction is in good agreement with the data for $(M/L)_{\rm cl} \lesssim 650\,h$ (M/L)$_\odot$, matching the number of satellites per host requires $(M/L)_{\rm cl} \gtrsim 800\,h$ (M/L)$_\odot$.

In order to investigate this failure of the models in more detail, and to explore all freedom in the CLF parameters, ideally one would construct a MGRS for each of the 2000 models in our MCMC. This, however, is not feasible computationally, and we therefore use an approximate, analytical method instead. The total number of galaxies in a flux-limited survey is given by

$$N_{\rm tot} = \int_0^\Omega {\rm d}\Omega \int_{z_{\rm min}}^{z_{\rm max}} {\rm d}z \frac{{\rm d}V}{{\rm d}\Omega\,{\rm d}z} \int_0^\infty {\rm d}M\, n(M)\, \langle N \rangle_{M,z}. \qquad (20)$$

Here Ω is the solid angle of sky of the survey, ${\rm d}V$ is the differential volume element, $z_{\rm min}$ and $z_{\rm max}$ are the survey redshift limits ($z_{\rm min} = 0.01$ and $z_{\rm max} = 0.15$ in our case), and $\langle N \rangle_{M,z}$ is the average number of galaxies per halo of mass M at redshift z which follows from the CLF according to

$$\langle N \rangle_{M,z} = \int_{L_{\rm min}(z)}^{L_{\rm max}(z)} \Phi(L|M)\,{\rm d}L \qquad (21)$$

with $L_{\rm min}(z)$ and $L_{\rm max}(z)$ the minimum and maximum luminosities of a galaxy at redshift z that makes the apparent magnitude limits of the survey (we adopt $15.0 < m_{b_J} < 19.3$ for the 2dFGRS). Similarly, the number of satellite galaxies follows from

$$N_{\rm sat} = \int_0^\Omega {\rm d}\Omega \int_{z_{\rm min}}^{z_{\rm max}} {\rm d}z \frac{{\rm d}V}{{\rm d}\Omega\,{\rm d}z} \int_0^\infty {\rm d}M\, n(M)\, \langle N_{\rm sat} \rangle_{M,z} \qquad (22)$$

with

$$\langle N_{\rm sat} \rangle_{M,z} = \begin{cases} \langle N \rangle_{M,z} - 1 & \text{if } \langle N \rangle_{M,z} \geqslant 1 \\ 0 & \text{otherwise.} \end{cases} \qquad (23)$$

The number of host galaxies in a flux-limited survey, finally, is given by

$$N_{\rm host} = \int_0^\Omega {\rm d}\Omega \int_{z_{\rm min}}^{z_{\rm max}} {\rm d}z \frac{{\rm d}V}{{\rm d}\Omega\,{\rm d}z} \int_0^\infty {\rm d}M\, n(M)\, w(M, z), \qquad (24)$$

where the weight function $w(M, z)$ is given by

$$w(M, z) = \begin{cases} 1 - \exp(-\langle N_{\rm sat} \rangle_{M,z}) & \text{if } \langle N \rangle_{M,z} \geqslant 1 \\ 0 & \text{otherwise} \end{cases}. \qquad (25)$$

This derives from the fact that only hosts with at least one satellite galaxy are counted as host galaxies, and using the fact that the number of satellite galaxies follows a Poisson distribution (see Section 3.2).

Using these equations we compute the abundances of host and satellite galaxies for each of the 2000 CLF models in the MCMC. However, this does not take care of incompleteness effects, fibre collisions or the size-projection effect. In addition, this assumes that all satellites are selected with zero interlopers. Instead, the host–satellite selection criterion used for the 2dFGRS and the MGRSs is not 100 per cent complete and yields about 15 per cent interlopers (van den Bosch et al. 2004b). To correct for all these effects/shortcomings we proceed as follows. We multiply the $N_{\rm host}$ and $N_{\rm sat}$ computed using equations (24) and (22) with correction factors $f_{\rm host}$ and $f_{\rm sat}$, respectively, which we calibrate using the results of the 12 MGRSs (that have been corrected for the size-projection effects) shown in Fig. 6 (open circles). We find a good match for $f_{\rm host} = 0.8$ and $f_{\rm sat} = 0.57 + 0.25[\log(M/L)_{\rm cl} - 3.0]$, the results of which are indicated by the solid line in the upper panels of Fig. 6. In what follows we assume that the same correction factors apply to all models in our MCMC.

Although only approximate, equations (20)–(25) combined with this simple scaling allows us to make predictions for the fractions of host and satellite galaxies for each of the 2000 CLF models in our MGRSs. The results are shown in the middle row of panels of Fig. 6. The overall behavior is similar to that of the 12 MGRSs shown in the upper panels: matching the fraction of host galaxies requires $(M/L)_{\rm cl} \lesssim 650\,h$ (M/L)$_\odot$, whereas matching the number of satellites per host basically requires the opposite. Meanwhile, the number of satellite galaxies is systematically too high. Although there are a few models that match the fraction of satellite galaxies of the 2dFGRS, as we show below, these models can not simultaneously match the fraction of host galaxies. Therefore, we conclude that *the abundances of host and satellite galaxies in the 2dFGRS cannot be reproduced within the ΛCDM concordance cosmology with $\sigma_8 = 0.9$.*

6.1 Evidence for a low value of σ_8

Using MGRSs similar to those presented here, Yang et al. (2004a) have shown that a ΛCDM concordance cosmology with $(M/L)_{\rm cl} = 500\,h$ (M/L)$_\odot$ predicts too much clustering power on small scales, and pairwise peculiar velocity dispersions that are too high. In Yang et al. (2004b) it was shown that the same model predicts too many large groups of galaxies. All these problems indicate that there are too many galaxies in (massive) clusters compared to observations. As demonstrated in Yang et al. (2004a,b), this can be remedied by adopting a high $(M/L)_{\rm cl}$ value, which results in fewer galaxies per cluster. This agrees perfectly with the fact that matching the number of satellites per host requires a similarly high $(M/L)_{\rm cl}$ value. However, the failure to match simultaneously the separate abundances of host and satellite galaxies, and the independent observational constraints on cluster mass-to-light ratios which indicate that $(M/L)_{\rm cl} \simeq (400 \pm 100)\,h$ (M/L)$_\odot$ (Carlberg et al. 1996; Fukugita et al. 1998; Bahcall et al. 2000), signal a clear shortcoming of this model.

An alternative solution to the apparent overabundance of cluster galaxies is to lower the abundance of cluster-sized haloes. This is most easily established by lowering the power-spectrum normalization σ_8. As shown by Yang et al. (2004a,b), a ΛCDM cosmology with $\sigma_8 = 0.7$ and $(M/L)_{\rm cl} = 500\,h$ (M/L)$_\odot$ can equally well match the clustering data, the pairwise peculiar velocities, and the abundances of galaxy groups as a concordance cosmology with $\sigma_8 = 0.9$ and $(M/L)_{\rm cl} = 900\,h$ (M/L)$_\odot$. In addition, as shown in van den Bosch et al. (2003b), this cosmology is in good agreement with the WMAP data and may even alleviate some problems with the ΛCDM cosmology regarding the concentration of dark matter haloes. It is interesting, therefore, to investigate whether such a model can also simultaneously match the abundances of host and satellite galaxies in the 2dFGRS.

In order to test this we construct a MCMC for the ΛCDM cosmology but with $\sigma_8 = 0.7$. We use the same number of steps and the same thinning factor as for the MCMC described in Section 2.2. The resulting distribution of parameters is indicated by the non-hatched histograms in Fig. 1. The corresponding median and 68 per cent confidence intervals are listed in Table 1 (model $\Lambda_{0.7}$) and correlations amongst various parameters are shown in Fig. 7. Compared to the distributions for the concordance cosmology with $\sigma_8 = 0.9$ the main differences are a reduction of the mean $(M/L)_{\rm cl}$, α_{15} and γ_3. Other parameters, notably M_1 and γ_1, are extremely insensitive to σ_8.

Unfortunately we do not have numerical simulations for a ΛCDM cosmology with $\sigma_8 = 0.7$, so that we cannot construct MGRSs for this cosmology (but see Yang et al. 2004a). Nevertheless, we can
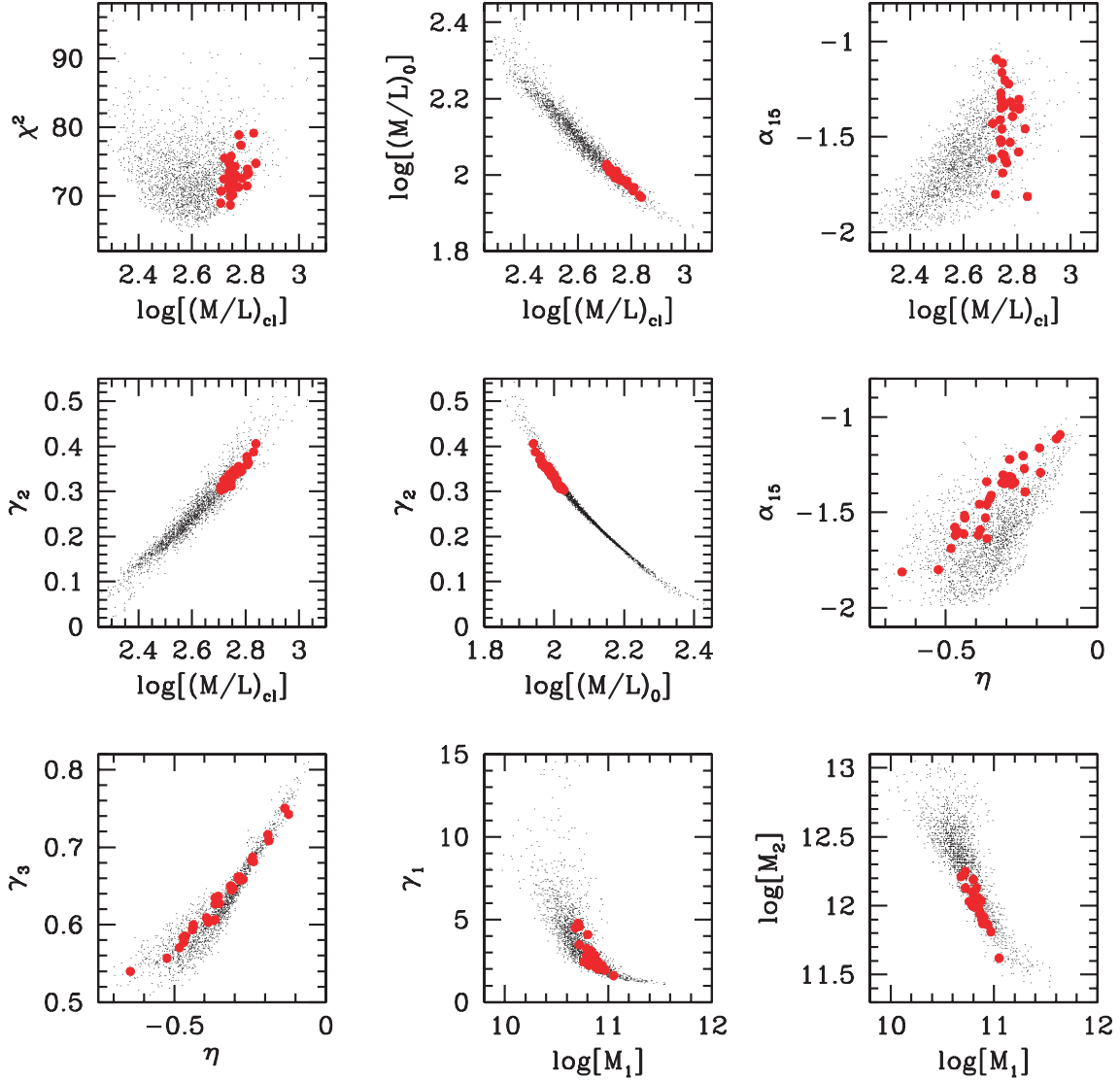
Figure 7. Same as Fig. 2 except that this time we plot parameter correlations for the MCMC constructed assuming a ΛCDM cosmology with $\sigma_8 = 0.7$. The thick solid dots indicate those models for which $\chi^2_{ab} < 4$, indicating a good match to the observed abundances of host and satellite galaxies. Note that these extra constraints severely restrict a number of the CLF parameters; cf. the parameters of models $\Lambda_{0.7}$ and $\Lambda^{ab}_{0.7}$ in Table 1.

use the scaling parameters f_{host} and f_{sat}, under the assumption that they are also valid for this cosmology, to predict the abundances of host and satellite galaxies for the low-σ_8 cosmology using equations (20)–(25). The resulting host–satellite fractions are shown in the lower panels of Fig. 6. Compared to the same results for the $\sigma_8 = 0.9$ cosmology (panels in middle row), the satellite fractions have been reduced, bringing them in better agreement with the 2dF-GRS results. In order to make the comparison between the two different σ_8 models more quantitative, and to investigate whether any model can simultaneously match the abundances of host and satellite galaxies, we introduce the goodness-of-fit measure

$$\chi^2_{ab} = \left(\frac{N_{host}/N_{tot} - 0.060}{0.002} \right)^2 + \left(\frac{N_{sat}/N_{tot} - 0.094}{0.003} \right)^2, \quad (26)$$

where the numbers in the numerators are the 2dFGRS ratios, and those in the denominators are the standard deviations owing to cosmic variance that we obtain from a set of independent MGRSs using

different simulation boxes (see Yang et al. 2004a).[4] For the $\sigma_8 = 0.7$ cosmology, the number of samples (out of a total of 2000) with $\chi^2_{ab} < (2, 4, 6)$ is (11, 34, 57). For the $\sigma_8 = 0.9$ cosmology, these numbers are (0, 0, 0). Clearly, *the $\sigma_8 = 0.7$ models are far more successful in simultaneously fitting the abundances of host and satellite galaxies than those with $\sigma_8 = 0.9$.*

The thick, solid dots in Fig. 7 indicate the 34 CLF models in the $\sigma_8 = 0.7$ cosmology for which $\chi^2_{ab} < 4$. Note how they are clustered together in parameter spaces. This means that using the observed abundances of host and satellite galaxies in the 2dFGRS as extra constraints allows a significant tightening of the constraints on the CLF parameters. We can take account of these additional constraints by weighting each sample in the MCMC with $\exp(-\chi^2_{ab}/2)$. This yields the median and 68 per cent confidence intervals listed under model $\Lambda^{ab}_{0.7}$ in Table 1. Note that the constraints on some of the CLF

---

[4] Note that these are very similar to the standard deviations obtained from a comparison of the host–satellite fractions in the 2dFGRS NGP and SGP.

parameters, notably $(M/L)_0$, $(M/L)_{cl}$ and M_1, are now much tighter (cf. model $\Lambda_{0.7}$).

Although the above results strongly favour a ΛCDM cosmology with relatively low power spectrum normalization, we caution that they are obtained assuming that the scaling parameters f_{host} and f_{sat}, determined for a small subset of all MCMC samples with $\sigma_8 = 0.9$, are valid for all MCMC samples, even those with $\sigma_8 = 0.7$. Testing the accuracy of this assumption requires similar *N*-body simulations as those used here, but for the ΛCDM cosmology with $\sigma_8 = 0.7$, and the construction of a large number of different MGRSs to probe the full CLF parameter space. Given our limited computational resources, we are unfortunately unable to perform these tests.

7 SUMMARY

Comparing the radial distribution and abundances of dark matter subhaloes in simulations with those of observed satellite galaxies, which are thought to be associated with these subhaloes, has led to two apparent inconsistencies: (i) The radial number density distribution of subhaloes reveals a constant density core, whereas galaxies in clusters seem to follow an NFW profile, and (ii) whereas dark matter haloes of different masses look homologous when it comes to the properties of their dark matter subhaloes, a galaxy cluster looks very different from a galaxy-sized system when it comes to their satellite galaxies (see Section 1 for references).

In this paper we have made use of the CLF to address these issues using data from the 2dFGRS. We constructed large MCMCs that allowed us to explore the full posterior distribution of the CLF parameter space. Using detailed MGRSs based on the CLF we have analysed the radial distribution and abundances of host and satellite galaxies in the 2dFGRS. Our main conclusions can be summarized as follows.

(i) The 2dFGRS is missing a significant fraction of galaxies in close (projected) pairs (about 2–4 per cent of *all* galaxies). This is in good agreement with Cole et al. (2001), who have shown that this close-pair deficiency is owing to the overlap and merging of galaxy images in the APM catalogue.

(ii) Owing to this close-pair incompleteness we cannot put strong constraints on the radial distribution of satellite galaxies. After modelling the close-pair incompleteness in our MGRSs the data are consistent with a model in which the radial number density distribution of satellite galaxies follows that of the dark matter particles (i.e. a NFW distribution), but we can not rule out alternatives with, for example, a constant number density core.

(iii) Within the ΛCDM concordance cosmology with $\sigma_8 = 0.9$ we can not simultaneously match the abundances of host and satellite galaxies in the 2dFGRS. Matching the number of satellites per host requires exceptionally high mass-to-light ratios on cluster scales. This is in excellent agreement with previous findings based on the CLF formalism, but severely overpredicts the abundance of both host and satellite galaxies.

(iv) Simultaneously matching the luminosity function, the clustering properties as a function of luminosity, and the abundances of host and satellite galaxies seems to require a reduction of the power spectrum normalization to $\sigma_8 \simeq 0.7$. As shown by Yang et al. (2004a,b), this is also required in order to match the pairwise peculiar velocities in the 2dFGRS, the clustering power on small scales, and the multiplicity function of galaxy groups.

(v) The CLF models for the ΛCDM concordance cosmology with $\sigma_8 = 0.7$ that match the abundances of host and satellite galaxies are extremely well constrained. They indicate that the average mass-to-light ratio of dark matter haloes reveals a pronounced minimum of 97 $\pm$ 3 h $(M/L)_{\odot}$ at a halo mass of $(7.2 \pm 0.7) \times 10^{10} \, h^{-1} \, M_{\odot}$. On cluster scales ($M \gtrsim 10^{14} \, h^{-1} \, M_{\odot}$) the average mass-to-light ratio is 570 $\pm$ 40 h $(M/L)_{\odot}$.

REFERENCES

Avila-Reese V., Firmani C., Klypin A., Kravtsov A. V., 1999, MNRAS, 310, 527
Bahcall N. A., Lubin L., Dorman V., 1995, ApJ, 447, L81
Bahcall N. A., Cen R., Davé R., Ostriker J. P., Yu Q., 2000, ApJ, 541, 1
Beers T. C., Tonry J. L., 1986, ApJ, 300, 557
Beijersbergen M., Hoekstra H., van Dokkum P. G., van der Hulst T., 2002, MNRAS, 329, 385
Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002, MNRAS, 333, 156
Berlind A. A. et al., 2003, ApJ, 593, 1
Brainerd T. G., Specian M. A., 2003, ApJ, 593, L7
Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517
Carlberg R. G., Yee H. K. C., Ellingson E., Abraham R., Gravel P., Morris S., Pritchet C. J., 1996, ApJ, 462, 32
Carlberg R. G., Yee H. K. C., Ellingson E., 1997, ApJ, 478, 462
Cole S., Kaiser N., 1989, MNRAS, 237, 1127
Cole S. et al. (The 2dFGRS Team), 2001, MNRAS, 326, 255
Colin P., Klypin A., Kravtsov A. V., Khokhlov M., 1999, ApJ, 523, 32
Colless M. et al. (the 2dFGRS team), 2001, MNRAS, 328, 1039
Courteau S., MacArthur L. A., Dekel A., van den Bosch F. C., McIntosh D. H., Dale D., 2003, preprint (astro-ph/0310440)
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
De Lucia G., Kauffmann G., Springel V., White S. D. M., Lanzoni B., Stoehr F., Tormen G., Yoshida N., 2004, MNRAS, 348, 333
Diaferio A., Kauffmann G., Michael B., White S. D. M., Schade D., Ellingson E., 2001, MNRAS, 323, 999
Diemand J., Moore B., Stadel J., 2004, MNRAS, 352, 535
D'onghia E., Lake G., 2004, ApJ, 612, 628
Efstathiou G., Bond J. R., White S. D. M., 1992, MNRAS, 285, 1
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 554, 114
Eke V. R. et al. (The 2dFGRS Team), 2004, MNRAS, 348, 866
Evans N. W., Wilkinson M. I., Guhathakurta P., Grebel E. K., Vogt S. S., 2000, ApJ, 540, L9
Fukugita M., Hogan C. J., Peebles P. J. E., 1998, ApJ, 503, 518
Gamerman D., 1997, Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. Chapman and Hall, London
Gao L., De Lucia G., White S. D. M., Jenkins A., 2004a, MNRAS, 352, L1
Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004b, MNRAS, 355, 819
Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 1998, MNRAS, 300, 146
Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 2000, ApJ, 544, 616
Gill S. P. D., Knebe A., Gibson B. K., Dopita M. A., 2004a, MNRAS, 351, 410
Gill S. P. D., Knebe A., Gibson B. K., 2004b, MNRAS, 351, 399
Hawkins E. et al. (The 2dFGRS Team), 2003, MNRAS, 346, 78
Hernández X., Avila-Reese V., Firmani C., 2001, MNRAS, 327, 329

Holmberg E., 1969, Ark. Astron., 5, 305

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Jing Y. P., 1998, ApJ, 503, L9

Jing Y. P., Suto Y., 2002, ApJ, 574, 538

Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201

Klypin A., Gottlöber S., Kravtsov A. V., Khokhlov A. M., 1999, ApJ, 516, 530

Klypin A., Zhao H., Somerville R. S., 2002, ApJ, 573, 597

Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004a, ApJ, 609, 35

Kravtsov A. V., Gnedin O. Y., Klypin A. A., 2004b, ApJ, 609, 482

Lake G., Tremaine S., 1980, ApJ, 238, L13

Lin Y.-T., Mohr J. J., Stanford S. A., 2004, ApJ, 610, 745

Little B., Tremaine S., 1987, 320, 493

Lorrimer S., Frenk C. S., Smith R. M., White S. D. M., Zaritsky D., 1994, MNRAS, 269, 696

Maddox S. J., Efstathiou G., Sutherland W. J., Loveday J., 1990, MNRAS, 243, 692

Madgwick D. S. et al., 2002, MNRAS, 333, 133

Magliocchetti M., Porciani C., 2003, MNRAS, 346, 186

Mateo M., 1998, ARA&A, 36, 435

McKay T. A. et al., 2002, ApJ, 571, L85

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347

Mo H. J., White S. D. M., 2002, MNRAS, 336, 112

Mo H. J., Yang X. H., van den Bosch F. C., Jing Y. P., 2004, MNRAS, 349, 205

Moore B., Ghigna S., Governato G., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493 (NFW)

Norberg P. et al. (the 2dFGRS team), 2002a, MNRAS, 332, 827

Norberg P. et al. (the 2dFGRS team), 2002b, MNRAS, 336, 907

Okamoto T., Habe A., 1999, ApJ, 516, 591

Prada F. et al., 2003, ApJ, 598, 260

Reed D., Governato F., Quinn T., Gardner J., Stadel J., Lake G., 2004, preprint (astro-ph/0406034)

Sales L., Lambas D. G., 2004, MNRAS, 348, 1236

Sandage A., Bingelli B., Tammann G. A., 1985, AJ, 90, 1759

Sanderson A. J. R., Ponman T. J., 2003, MNRAS, 345, 1241

Scranton R., 2003, MNRAS, 339, 410

Seljak U., 2000, MNRAS, 318, 203

Seljak U., Warren M. S., 2004, MNRAS, 355, 129

Sheth R. K., Tormen, G., 1999, MNRAS, 308, 119

Sheth R. K., Tormen G., 2002, MNRAS, 329, 61

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Smith R. E. et al., 2003, MNRAS, 341, 1311

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Stoehr F., White S. D. M., Tormen G., Springel V., 2002, MNRAS, 335, L84

Tormen G., 1997, MNRAS, 290, 411

Trentham N., Hodgkin S., 2002, MNRAS, 333, 423

Trentham N., Tully R. B., 2002, MNRAS, 335, 712

Vader J. P., Sandage A., 1991, ApJ, 79, L1

Vale A., Ostriker J. P., 2004, MNRAS, 253, 189

van den Bosch F. C., 2002, MNRAS, 332, 456

van den Bosch F. C., Yang X. H., Mo H. J., 2003a, MNRAS, 340, 771

van den Bosch F. C., Mo H. J., Yang X. H., 2003b, MNRAS, 345, 923

van den Bosch F. C., Yang X. H., Mo H. J., 2004a, MNRAS, 348, 735

van den Bosch F. C., Norberg P., Mo H. J., Yang X. H., 2004b, MNRAS, 352, 1302

van der Marel R. P., Magorrian J., Carlberg R. G., Yee H. K. C., Ellingson E., 2000, AJ, 119, 2038

Wang Y., Yang X. H., Mo H. J., van den Bosch F. C., Chu Y., 2004, MNRAS, 353, 287

Weller J., Ostriker J. P., Bode P., 2004, preprint (astro-ph/0405445)

White M., 2002, ApJS, 143, 241

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

Willman B., Governato F., Dalcanton J. J., Reed D., Quinn T., 2004, MNRAS, 353, 639

Yan R., Madgwick D. S., White M., 2003, ApJ, 598, 848

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

Yang X. H., Mo H. J., Jing Y. P., van den Bosch F. C., Chu Y., 2004a, MNRAS, 350, 1153

Yang X. H., Mo H. J., van den Bosch F. C., Jing Y. P., 2004b, MNRAS, in press (doi:10.1111/j.1365-2966.2005.08560)

Zaritsky D., Gonzales A. H., 1999, PASP, 111, 1508

Zaritsky D., White S. D. M., 1994, ApJ, 435, 599

Zaritsky D., Smith R., Frenk C. S., White S. D. M., 1993, ApJ, 405, 464

Zaritsky D., Smith R., Frenk C. S., White S. D. M., 1997, ApJ, 478, 39

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