GALAXY GROUPS IN THE SDSS DR4. I. THE CATALOG AND BASIC PROPERTIES

XIAOHU YANG,[1,2] H. J. MO,[3] FRANK C. VAN DEN BOSCH,[4] ANNA PASQUALI,[4] CHENG LI,[1,2] AND MARCO BARDEN[4,5]

ABSTRACT

We use a modified version of the halo-based group finder developed by Yang et al. to select galaxy groups from the Sloan Digital Sky Survey (SDSS DR4). In the first step, a combination of two methods is used to identify the centers of potential groups and to estimate their characteristic luminosity. Using an iterative approach, the adaptive group finder then uses the average mass-to-light ratios of groups, obtained from the previous iteration, to assign a tentative mass to each group. This mass is then used to estimate the size and velocity dispersion of the underlying halo that hosts the group, which in turn is used to determine group membership in redshift space. Finally, each individual group is assigned two different halo masses: one based on its characteristic luminosity and the other based on its characteristic stellar mass. Applying the group finder to the SDSS DR4, we obtain 301,237 groups in a broad dynamic range, including systems of isolated galaxies. We use detailed mock galaxy catalogs constructed for the SDSS DR4 to test the performance of our group finder in terms of completeness of true members, contamination by interlopers, and accuracy of the assigned masses. This paper is the first in a series and focuses on the selection procedure, tests of the reliability of the group finder, and the basic properties of the group catalog (e.g., the mass-to-light ratios, the halo mass–to–stellar mass ratios). The group catalogs including the membership of the groups are available on request.

Subject headings: dark matter — galaxies: halos — large-scale structure of universe — methods: statistical

Online material: color figures

1. INTRODUCTION

Galaxies are thought to form and reside in extended cold dark matter halos. One of the ultimate challenges in astrophysics is therefore to obtain a detailed understanding of how galaxies with different physical properties occupy dark matter halos of different mass. This relationship not only conveys important information about how different galaxies form and evolve in different dark matter halos, but it also provides the necessary basis for translating the observed distribution of galaxies into the large-scale distribution of matter throughout the universe.

Theoretically, the relationship between galaxies and dark matter halos can be studied using numerical simulations (e.g., Katz et al. 1996; Pearce et al. 2000; Springel 2005; Springel et al. 2005) or semianalytical models (e.g., White & Frenk 1991; Kauffmann et al. 1993, 2004; Somerville & Primack 1999; Cole et al. 2000; van den Bosch 2002; Kang et al. 2005; Croton et al. 2006). Both of these techniques try to model the process of galaxy formation ab initio. However, since our understanding of the various physical processes involved is still relatively poor, the relations between the properties of galaxies and their dark matter halos predicted by these simulations and semianalytical models still need to be tested against observations.

More recently, the halo occupation model has opened another avenue to probe the galaxy–dark matter connection (e.g., Jing et al. 1998; Peacock & Smith 2000; Berlind & Weinberg 2002; Cooray & Sheth 2002; Scranton 2003; Yang et al. 2003; van den Bosch et al. 2003, 2007; Yan et al. 2003; Tinker et al. 2005; Zheng et al. 2005; Cooray 2006; Vale & Ostriker 2006). This technique uses the observed galaxy luminosity function and two-point correlation functions to constrain the average number of galaxies of given properties that occupy a dark matter halo of given mass. Although this method has the advantage that it can typically yield much better fits to the data than the semianalytical models or numerical simulations, one typically needs to assume a somewhat ad hoc functional form to describe the halo occupation model.

A more direct way of studying the galaxy-halo connection is by using galaxy groups, provided that these are defined as sets of galaxies that reside in the same dark matter halo.[6] With a well-defined galaxy group catalog, one cannot only study the properties of galaxies as a function of their group properties (e.g., Yang et al. 2005a, 2005d; Collister & Lahav 2005; van den Bosch et al. 2005; Robotham et al. 2006; Zandivarez et al. 2006; Weinmann et al. 2006a, 2006b) but one can also probe how dark matter halos trace the large-scale structure of the universe (e.g., Yang et al. 2005c, 2006; Coil et al. 2006; Berlind et al. 2006). During the past two decades, numerous group catalogs have been constructed from various galaxy redshift surveys, most notably the Center for Astrophysics Redshift Survey (e.g., Geller & Huchra 1983), the Las Campanas Redshift Survey (e.g., Tucker et al. 2000), the Two Degree Field Galaxy Redshift Survey (2dFGRS; Merchán & Zandivarez 2002; Eke et al. 2004; Yang et al. 2005b; Tago et al. 2006; Einasto et al. 2007), the high-redshift DEEP2 survey (Gerke et al. 2005), and the Two Micron All Sky Redshift Survey (Crook et al. 2007). Various group catalogs have also been constructed from the redshift samples selected from the ongoing Sloan Digital Sky Survey (SDSS): Goto (2005) and Berlind et al. (2006) used a friends-of-friends (FOF) algorithm to identify groups in the SDSS Data Release 2 (DR2; Abazajian et al. 2004),

[1] Shanghai Astronomical Observatory, The Partner Group of MPA, Nandan Road 80, Shanghai 200030, China; xhyang@shao.ac.cn.

[2] Joint Institute for Galaxy and Cosmology (JOINGC) of Shanghai Astronomical Observatory and University of Science and Technology of China.

[3] Department of Astronomy, University of Massachusetts, Amherst, MA 01003-9305.

[4] Max-Planck-Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany.

[5] Institut für Astrophysik, Leopold-Franzens Universität Innsbruck, Technikerstrasse 25, A-6020 Innsbruck, Austria.

[6] In this paper we refer to a system of galaxies as a group regardless of its richness, including isolated galaxies (i.e., groups with a single member) and rich clusters of galaxies.

Miller et al. (2005) used the C4 algorithm to find clusters in the SDSS DR2, Weinmann et al. (2006a) used the halo-based group finder of Yang et al. (2005b) to identify groups in the New York University Value-Added Galaxy Catalog (NYU-VAGC) of Blanton et al. (2005) which is also based on the SDSS DR2, and Merchán & Zandivarez (2005) used a FOF algorithm to identify groups in the SDSS DR3 (Abazajian et al. 2005). Group catalogs have also been constructed from the SDSS photometric data. Goto et al. (2002) developed a cut-and-enhance method and applied it to the early SDSS commissioning data. Bahcall et al. (2003) compared the properties of groups selected from the early SDSS commissioning data with two different selection methods, a hybrid matched filter method (Kim et al. 2002) and a "maxBCG" method developed by Annis et al. (1999). Lee et al. (2004) identified compact groups in the SDSS Early Data Release (EDR; Stoughton et al. 2002). More recently, Koester et al. (2007) used the "maxBCG" method to assemble a large photometrically selected galaxy group catalog from the SDSS with a sky coverage of $\sim$7500 deg^2. Photometric catalogs also exist outside the SDSS (e.g., Gonzalez et al. 2001; Gladders & Yee 2005).

In a recent paper, Yang et al. (2005b) developed a halo-based group finder that is optimized for grouping galaxies that reside in the same dark matter halo. Using mock galaxy redshift surveys constructed from the conditional luminosity function model (see Yang et al. 2004), they found that this group finder is very successful in associating galaxies according to their common dark matter halos. In particular, the group finder also performs reliably for poor systems, including isolated galaxies in small mass halos. This makes this halo-based group finder ideally suited to study the relation between galaxies and dark matter halos over a wide dynamic range in halo masses. Thus far, the halo-based group finder has been applied to both the 2dFGRS (Yang et al. 2005b) and to the SDSS DR2 (Weinmann et al. 2006a). In this paper we apply a slightly modified and improved version to the NYU-VAGC based on the SDSS DR4. As the first in a series, this paper focuses on the selection process and the basic properties of the group catalog. More detailed analyses of the group properties and the implications for halo occupation statistics and galaxy formation will be presented in forthcoming papers.

This paper is organized as follows. Section 2 gives a brief description of the SDSS data used in this paper. In § 3 we describe the halo-based group finder and the methods to assign halo masses to the groups. In § 4 we present the group catalog based on the SDSS DR4 and study some of its basic properties. Finally, we summarize our results in § 5. Unless stated otherwise, we adopt a ΛCDM cosmology with parameters that are consistent with the three-year data release of the WMAP mission (hereafter WMAP3 cosmology): $\Omega_m = 0.238$, $\Omega_\Lambda = 0.762$, $\Omega_b = 0.042$, $n = 0.951$, $h = H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1}) = 0.73$, and $\sigma_8 = 0.75$ (Spergel et al. 2007).

2. GALAXY SAMPLES

The data used in this paper is taken from the Sloan Digital Sky Survey (SDSS; York et al. 2000), a joint five-passband (u, g, r, i, and z) imaging and medium-resolution ($R \sim 1800$) spectroscopic survey. More specifically, we make use of the New York University Value-Added Galaxy Catalog (NYU-VAGC; see Blanton et al. 2005), which is based on SDSS DR4 (Adelman-McCarthy et al. 2006) but includes a set of significant improvements over the original pipelines. From this catalog we select all galaxies in the main galaxy sample with redshifts in the range $0.01 \leq z \leq 0.20$ and with a redshift completeness $\mathcal{C} > 0.7$ (about 4% of the galaxies have $\mathcal{C} \leq 0.7$). This leaves a grand total of 362,356 gal-

axies with reliable r-band magnitudes and with measured redshifts from the SDSS. We refer to this sample of galaxies as sample I.

In addition, there are 7091 galaxies with $0.01 \leq z \leq 0.20$ in the NYU-VAGC which have redshifts from alternative sources: from the 2dFGRS (Colless et al. 2001), from the PSCz (Saunders et al. 2000), or from the RC3 (de Vaucouleurs et al. 1991).[7] Including these galaxies results in sample II, with a total of 369,447 galaxies. As an illustration, Figure 1 shows the sky coverage ($\sim$4514 deg^2) of all galaxies in sample II in Galactic coordinates, overlaid on the galactic extinction contours of Schlegel et al. (1998).

The two samples described above suffer from incompleteness due to fiber collisions. No two fibers on the same SDSS plate can be closer than 55″. Although this fiber-collision constraint is partially alleviated by the fact that neighboring plates have overlap regions, $\sim$7% of all galaxies eligible for spectroscopy do not have a measured redshift. Hereafter, we refer to these galaxies as "fiber-collision" galaxies. Since fiber collisions are more frequent in regions of high (projected) density, they are more likely to occur in richer groups, thus causing a systematic bias that may need to be accounted for. A simple method of doing so is to assign a galaxy which lacks an observed redshift due to fiber collisions the redshift of the galaxy with which it collided. As shown in Zehavi et al. (2002), roughly 60% of the fiber-collision galaxies have a redshift within 500 km s^{-1} of their nearest neighbor, and for these cases, the above procedure is more than appropriate. However, there are also cases in which the fiber-collision galaxy has a true redshift that is very different from that of its nearest neighbor. If the fiber-collision galaxy is assigned a redshift that is too large, its implied luminosity will also be too large and can in fact become excessively large. This in turn can have dramatic consequences for our group finder. To limit the impact of these catastrophic failures we remove the $\sim$1.0% of all fiber-collision galaxies that have an implied absolute magnitude of $^{0.1}M_r - 5 \log h \leq -22.5$ (see eq. [1] below). In our redshift interval, there are a total of 38,672 galaxies with an assigned redshift and with $^{0.1}M_r - 5 \log h > -22.5$. Including these galaxies results in sample III, with a total of 408,119 galaxies.

In what follows, we use sample II as our main sample for selecting galaxy groups. For completeness, we also apply our group finder over samples I and III, and we occasionally compare results based on all three group catalogs.

2.1. Magnitudes and Stellar Masses

For each galaxy we compute the absolute magnitude in bandpass Q using

$$^{0.1}M_Q - 5 \log h = m_Q + \Delta m_Q - \text{DM}(z) - K_Q - E_Q, \quad (1)$$

where $\text{DM}(z) = 5 \log[D_L/(h^{-1} \text{ Mpc})] + 25$ is the bolometric distance modulus calculated from the luminosity distance D_L using a WMAP3 cosmology with $\Omega_m = 0.24$ and $\Omega_\Lambda = 0.76$. Δm_Q is the latest zero-point correction for the apparent magnitudes, which converts the SDSS magnitudes to the AB system, and for which we adopt $\Delta m_Q = (-0.036, +0.012, +0.010, +0.028, +0.040)$ for $Q = (u, g, r, i, z)$ (M. Blanton 2006, private communication). All absolute magnitudes are $K + E$ corrected to $z = 0.1$. For the K corrections we use the latest version of "Kcorrect" (v4) described in Blanton et al. (2003a; see also Blanton & Roweis 2007), which we apply to *all* galaxies that have meaningful magnitudes and meaningful redshifts, including those that have redshifts from
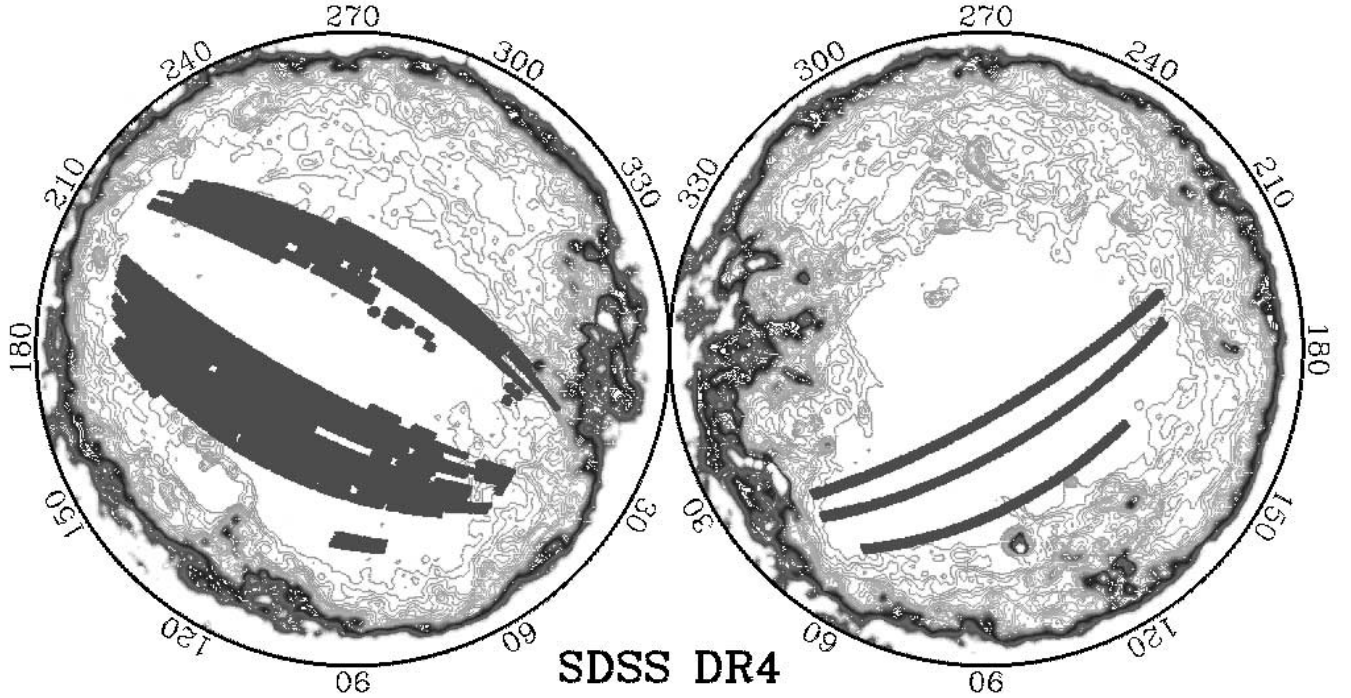
---

[7] See Blanton et al. (2005) for details.

Fig. 1.—Sky coverage of the SDSS DR4 galaxies in sample II, overlaid on the galactic extinction contours of Schlegel et al. (1998). Note that the SDSS probes regions of low galactic extinction. [*See the electronic edition of the Journal for a color version of this figure.*]

alternative sources and those that have been assigned the redshift of their nearest neighbor. Finally, the evolution corrections to $z = 0.1$ are computed using $E_Q = A_Q(z - 0.1)$, with $A_Q = (-4.22, -2.04, -1.62, -1.61, -0.76)$ for $Q = (u, g, r, i, z)$ (see Blanton et al. 2003a). Note that these evolution corrections imply that galaxies were brighter in the past (at higher redshifts).

In addition to the absolute magnitudes, we also compute for each galaxy its stellar mass, M_*. Using the relation between the stellar mass–to-light ratio and color of Bell et al. (2003), we obtain

$$\log\left(\frac{M_*}{h^{-2}\ M_\odot}\right) = -0.306 + 1.097[^{0.0}(g - r)] - 0.1$$
$$- 0.4\left(^{0.0}M_r - 5\log h - 4.64\right), \quad (2)$$

where $^{0.0}(g - r)$ and $^{0.0}M_r - 5\log h$ are the $(g - r)$ color and r-band magnitude $K + E$ corrected to $z = 0.0$, 4.64 is the r-band magnitude of the Sun in the AB system (Blanton & Roweis 2007), and the -0.10 term effectively implies that we adopt a Kroupa (2001) IMF (Borch et al. 2006).

For a small fraction of all galaxies, the $g - r$ color that results from the photometric SDSS pipeline is unreliable. These galaxies typically have $g - r$ colors that are clearly unrealistic (they are catastrophic outliers in the color-magnitude distribution). If this is not accounted for, equation (2) assigns these galaxies stellar masses that are unrealistically high or low, which can have a dramatic impact on our group finder (which assigns masses to the groups based on their characteristic stellar mass; see § 3.5 below). To take account of these outliers we proceed as follows. As shown by Baldry et al. (2004) the distribution of $(g - r)$ colors at a given r-band magnitude can be well approximated by a bi-Gaussian function, representing the red sequence and the blue

cloud. Following Li et al. (2006) we therefore fit bi-Gaussian functions to the distribution of $^{0.0}(g - r)$ for a total of 118 bins in $^{0.0}M_r - 5\log h$. As shown in Li et al. (2006), these fits accurately capture the distribution of galaxies in the color-magnitude plane. For any galaxy that falls outside the 3σ ranges from the mean color-magnitude relations of both the red sequence and the blue cloud ($\sim$2% of all galaxies in sample III), we compute its stellar mass using the mean color of the red sequence (when the galaxy is too red) or the blue cloud (when the galaxy is too blue). Detailed tests have shown that this prevents any problems with catastrophic outliers.

3. THE CONSTRUCTION OF THE GROUP CATALOG

3.1. *The Group Finder*

The group finder adopted here is similar to that developed in Yang et al. (2005b). The strength of this group finder, hereafter referred to as the halo-based group finder, is that it is iterative and based on an adaptive filter modeled after the general properties of dark matter halos. In addition, unlike the traditional FOF method, this group finder can also identify groups with only a single member, which allows us to sample a wider dynamic range in group masses. Note that various masses are used in our group finder and in the presentation. In order to avoid confusion, we list in Table 1 the various masses that are used along with their definitions.

The halo-based group finder consists of the following main steps:

1. *Find potential group centers.*—We use a combination of two different methods to identify the centers (and members) of potential groups in redshift space. First, we use the traditional FOF algorithm (e.g., Davis et al. 1985) with very small linking lengths in redshift space to assign galaxies into tentative groups that may

Name	Definition
M_*	Stellar mass of a galaxy
$M_{stellar}$	Total stellar mass of group members with $^{0.1}M_r - 5 \log h \leq -19.5$
M_h	True halo mass (unless stated otherwise)
M_L	Halo mass estimated using the ranking of $L_{19.5}$
M_S	Halo mass estimated using the ranking of $M_{stellar}$

represent the central parts of groups. The linking lengths adopted are $\ell_z = 0.3$ along the line of sight and $\ell_p = 0.05$ in the transverse direction, both in units of the mean separation of galaxies at the redshift in question. The geometrical, luminosity-weighted centers of all the FOF groups thus identified with two members or more are considered as the centers of potential groups. Next, for all galaxies not yet linked to these FOF groups, we treat them also as tentative centers of potential groups.

2. *Determine the characteristic luminosity of each tentative group.*—In order to be able to meaningfully compare different groups, we define the group's *characteristic* luminosity, $L_{19.5}$, defined as the combined luminosity of all group members with $^{0.1}M_r - 5 \log h \leq -19.5$ (here again, all absolute magnitudes are $K + E$ corrected to $z = 0.1$). For groups with redshifts $z < 0.09$, all galaxies with $^{0.1}M_r - 5 \log h \leq -19.5$ make the flux limit of the SDSS spectroscopic sample, and $L_{19.5}$ can be computed directly using

$$L_{19.5} = \sum_i L_i / \mathcal{C}_i, \qquad (3)$$

where L_i is the luminosity of the ith member galaxy, $\mathcal{C}_i$ is the completeness of the survey at the position of that galaxy, and the summation is over all group members with $^{0.1}M_r - 5 \log h \leq -19.5$. For groups with $z > 0.09$, however, we need to correct for the missing members with $^{0.1}M_{r,\lim} - 5 \log h \leq^{0.1} M_r - 5 \log h \leq -19.5$, with $^{0.1}M_{r,\lim} - 5 \log h$ the absolute magnitude limit at the redshift of the group. In this case, we define the characteristic luminosity as

$$L_{19.5} = \frac{1}{f(L_{19.5}, L_{\lim})} \sum_i \frac{L_i}{\mathcal{C}_i}, \qquad (4)$$

with $f(L_{19.5}, L_{\lim})$ a correction factor $(0 < f \leq 1)$ that accounts for the galaxies missed because of the apparent magnitude limit of the spectroscopic survey. The method of computing $f(L_{19.5}, L_{\lim})$ is described in § 3.3 below.

3. *Estimate the mass, size, and velocity dispersion of the dark matter halo associated with each tentative group.* —Using the value of $L_{19.5}$ determined above and an assumption for the group mass-to-light ratio, $M_h/L_{19.5}$, we assign each tentative group with a halo mass which we use in the following steps to assign group memberships.

In the first iteration we simply adopt a constant mass-to-light ratio, $M_h/L_{19.5} = 500 \, h \, M_\odot/L_\odot$ for all groups. For all subsequent iterations, however, we use the $M_h/L_{19.5}$-$L_{19.5}$ relation obtained from the previous iteration (using the method described in § 3.5). Because of this iterative technique, the final group catalog is very insensitive to the (fairly arbitrary) initial guess of $M_h/L_{19.5} = 500 \, h \, M_\odot/L_\odot$ (see Yang et al. 2005b for detailed tests). Note that the halo masses in this step are estimated using the mass-to-light ratio and agree well with the final masses to be estimated in § 3.5.

Throughout this paper we define dark matter halos as having an overdensity of 180. This implies, for the WMAP3 cosmology adopted here, a halo radius of

$$r_{180} = 1.26 \, h^{-1} \, \text{Mpc} \left(\frac{M_h}{10^{14} \, h^{-1} \, M_\odot} \right)^{1/3} (1 + z_{group})^{-1}, \quad (5)$$

where z_{group} is the redshift of the group center, and a line-of-sight velocity dispersion of

$$\sigma = 397.9 \, \text{km s}^{-1} \left(\frac{M_h}{10^{14} \, h^{-1} \, M_\odot} \right)^{0.3214}. \qquad (6)$$

The latter is a fitting function that accurately captures the halo mass dependence of the one-dimensional velocity dispersion as given by equation (14) in van den Bosch et al. (2004), using the halo concentrations of Maccio et al. (2007).

4. *Update group memberships using halo information.*—Once we have a group center and a tentative estimate of the size, mass, and velocity dispersion of the halo associated with it, we can assign galaxies to this group using these halo properties. If we assume that the distribution of galaxies in phase space follows that of the dark matter particles, the number density contrast of galaxies in the redshift space around the group center (assumed to coincide with the center of the halo) at redshift z_{group} can be written as

$$P_M(R, \Delta z) = \frac{H_0}{c} \frac{\Sigma(R)}{\bar{\rho}} p(\Delta z), \qquad (7)$$

where c is the speed of light, $\Delta z = z - z_{group}$, $\bar{\rho}$ is the average density of universe, and $\Sigma(R)$ is the projected surface density of a (spherical) NFW (Navarro et al. 1997) halo

$$\Sigma(R) = 2 r_s \bar{\delta} \bar{\rho} f(R/r_s), \qquad (8)$$

with r_s the scale radius,

$$f(x) = \begin{cases} \dfrac{1}{x^2 - 1} \left\{ 1 - \dfrac{\ln\left[\left(1 + \sqrt{1 - x^2}\right)/x\right]}{\sqrt{1 - x^2}} \right\}, & x < 1, \\[3mm] 1/3, & x = 1, \\[3mm] \dfrac{1}{x^2 - 1} \left(1 - \dfrac{\arctan\sqrt{x^2 - 1}}{\sqrt{x^2 - 1}} \right), & x > 1, \end{cases} \qquad (9)$$

$$\bar{\delta} = \frac{180}{3} \frac{c_{180}^3}{\ln(1 + c_{180}) - c_{180}/(1 + c_{180})}, \qquad (10)$$

with $c_{180} = r_{180}/r_s$. The function $p(\Delta z) \, d\Delta z$ describes the redshift distribution of galaxies within the halo and is assumed to have a Gaussian form,

$$p(\Delta z) = \frac{1}{\sqrt{2\pi}} \frac{c}{\sigma(1 + z_{group})} \exp\left[\frac{-(c\Delta z)^2}{2\sigma^2(1 + z_{group})^2} \right], \quad (11)$$

where σ is the rest-frame velocity dispersion of equation (6). Thus defined, $P_M(R, \Delta z)$ is the three-dimensional density contrast in redshift space. In order to decide whether a galaxy should be assigned to a particular group we proceed as follows. For each galaxy we loop over all groups and compute the distance $(R, \Delta z)$ between the galaxy and the group center, where R is the projected distance at the redshift of the group. If $P_M(R, \Delta z) \geq B$, with

$B = 10$ an appropriately chosen background level (see Yang et al. 2005b), the galaxy is assigned to the group. If a galaxy can be assigned to more than one group according to this criterion, it is only assigned to the one for which $P_M(R, \Delta z)$ is the largest. Finally, if all members in two groups can be assigned to one group according to the above criterion, the two groups are merged into a single group.

5. *Iterate.*—Using the new memberships obtained in step 4, we recompute the group centers and go back to step 2. This iterating process goes on until there is no further change in the group memberships. Next we use the resulting group catalog to compute $f(L_{19.5}, L_{\rm lim})$ and the relation between $M_h/L_{19.5}$ and $L_{19.5}$, and we go back to step 1. We stop this iteration cycle once the $M_h/L_{19.5}$-$L_{19.5}$ relation has converged, which typically takes only three to four iterations.

3.2. *Completeness, Contamination, and Purity of the Group Catalogs*

To test its performance, we run our halo-based group finder over a detailed mock galaxy redshift survey (MGRS) that mimics the SDSS DR4. The MGRS is constructed by populating dark matter halos in numerical simulations of cosmological volumes with galaxies of different luminosities, using the conditional luminosity function (CLF) model of van den Bosch et al. (2007, in preparation). This CLF describes the halo occupation statistics of SDSS galaxies and accurately matches the SDSS luminosity function and the clustering properties of SDSS galaxies as a function of their luminosity. We used a stack of simulations with different resolutions (100 and 300 h^{-1} Mpc cubes with 512^3 dark matter particles each) to make sure that the mock catalog is complete down to the SDSS magnitude limit (see Yang et al. 2004 for the stacking). Next, a MGRS is constructed mimicking the sky coverage of the SDSS DR4 and taking detailed account of the angular variations in the magnitude limits and completeness of the data (see Li et al. 2007 for details). Methods like this are becoming widespread for both understanding cluster detection (Yang et al. 2005b; Gerke et al. 2005; Koester et al. 2007; Cohn et al. 2007) and in quantifying selection functions (Rozo et al. 2007).

To assess the performance of the group finder we follow Yang et al. (2005b) and proceed as follows. For each group k, we look up the halo ID, h_k, of the brightest group member, and we define N_t as the total number of true members in the MGRS (with $0.01 \leq z \leq 0.20$) that belong to halo h_k, N_s as the number of these true members that are selected as members of group k, N_i as the number of interlopers (group members that belong to a different halo), and N_g as the total number of selected group members. These allow us to define, for each group, the following three quantities:

Completeness.—$f_c \equiv N_s/N_t$,
Contamination.—$f_i \equiv N_i/N_t$,
Purity.—$f_p \equiv N_t/N_g$.

Since $N_g = N_i + N_s$, we have that $f_p = 1/(f_c + f_i)$. A purity $f_p < 1$ implies that the number of interlopers is larger than the number of missed true members, while $f_p > 1$ implies that the group is not complete ($f_c < 1$) and the number of missed true members is larger than the number of interlopers. Note that the identity of the halo that belongs to a group is solely based on the halo ID of the brightest group member. Consequently, the contamination f_i can be larger than unity. An ideal, perfect group finder yields groups with $f_c = f_p = 1$ and $f_i = 0$. In the case of the halo-based group finder used here, the value for the background level B has been tuned to maximize the average value of $f_c(1 - f_i)$ (see Yang et al. 2005b).

Results obtained from the MGRS are shown in Figure 2. Since groups with a single member have zero contamination ($f_i = 0$) by definition, results are only shown for groups with a richness $N \geq 2$. The top left panel of Figure 2 shows the cumulative distributions of the completeness f_c. Different line styles correspond to groups of different true halo masses, as indicated. The fraction of groups with 100% completeness (i.e., with $f_c = 1$) depends on group mass and ranges from ~95% for low-mass groups to ~60% for the most massive clusters. Since our group finder is tuned to maximize the average value of $f_c(1 - f_i)$, massive groups with larger velocity dispersions have larger f_i due to the contamination of foreground and background galaxies. A compromise between f_c and f_i leads to smaller f_c for more massive groups. Almost independent of group mass, we find that more than 90% of all groups have a completeness $f_c > 0.6$, while an average of 80% of all groups have $f_c > 0.8$. The middle left panel of Figure 2 shows the cumulative distributions of the contamination f_i. On average, around 65% of the groups have zero contamination ($f_i = 0$), while ~85% of the groups have $f_i \leq 0.5$, again virtually independent of group mass. These interlopers (contamination) are either nearby field galaxies or the member galaxies of nearby massive groups, especially those along the line of sight. Finally, the bottom left panel of Figure 2 shows the cumulative distributions of the purity f_p, indicating that there are on average as many groups with $f_p < 1$ as with $f_p > 1$. The break at $f_p \sim 1$ means that the number of recovered group members is about the same as the number of true members. Thus, the sharper the break is, the better. An ideal situation is a step function at $f_p = 1$. In addition, only a negligibly small fraction of groups have $f_p < 0.5$, while only for the most massive halos is there a significant fraction (~10%) with $f_p > 1.5$.

We also determine the completeness, contamination, and purity in terms of the total luminosity rather than the number of member galaxies. The corresponding results are shown in the right panels of Figure 2, from top to bottom. As one can see, the results are very similar to those in terms of the number of members.

As a final, quantitative assessment of the performance of our halo-based group finder, we examine the *global completeness*, $f_{\rm halo}$, defined as the fraction of halos in the MGRS whose brightest member has actually been identified as the brightest (central) galaxy of its group. Figure 3 shows $f_{\rm halo}$ obtained from our CLF mock for halos with $N_t \geq 1$ (*dashed lines*) and $N_t \geq 3$ (*solid lines*) as functions of the true halo mass. As one can see, the group finder successfully selects more than 90% of all the true halos with masses $\gtrsim 10^{12}$ h^{-1} $M_\odot$ almost independent of their richness and with only a very weak dependence on halo mass. Note that this does not imply that ~10% of the central galaxies in dark matter halos have not been selected by the group finder. Especially for the more massive halos, the vast majority of these central galaxies have been selected as a group member, but they are not the brightest group member. This can happen whenever two nearby halos are merged into a single group by the group finder.

3.3. *Completeness Corrections for the Characteristic Luminosity and Stellar Mass*

An important parameter for each group is its characteristic luminosity, defined by equation (4). Since the correction factor $f(L_{19.5}, L_{\rm min})$ depends on the characteristic luminosity $L_{19.5}$ itself, it can only be determined in an iterative way. In the first iteration of our group finder, we use

$$f(L_{19.5}, L_{\rm lim}) = \frac{\int_{L_{\rm lim}}^{\infty} L\phi(L)\,dL}{\int_{L_{\rm cut}}^{\infty} L\phi(L)\,dL}, \qquad (12)$$
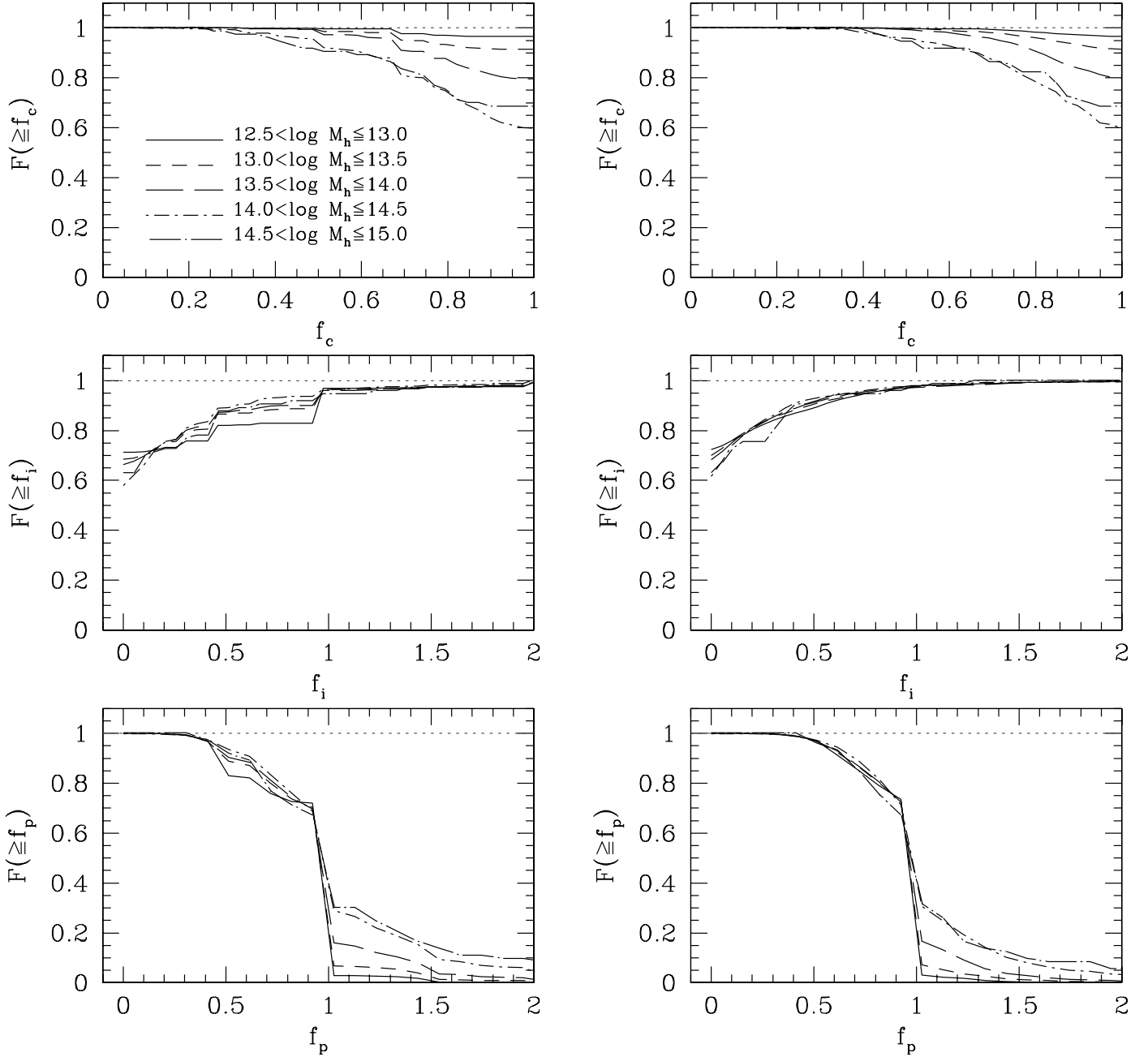
Fig. 2.—Cumulative distributions of completeness, f_c (the fraction of true members; *top*), contamination, f_i (the fraction of interlopers; *middle*), and purity, f_p (ratio between the true members and the group members; *bottom*). See text for the detailed definitions of all three parameters. In the left and right panels, these values are number weighted and luminosity weighted, respectively. Different lines show the results for groups in halos of different masses, as indicated. Results are plotted for groups with at least two members, since groups with only one member have, by definition, $f_i = 0$. [*See the electronic edition of the Journal for a color version of this figure.*]

where L_{cut} is the luminosity that corresponds to $^{0.1}M_r - 5 \log h = -19.5$, and $\phi(L)$ is the galaxy luminosity function, here assumed to be that obtained by Blanton et al. (2003b). However, as discussed in Yang et al. (2005b), it is not reliable to make the correction based on the assumption that the galaxy luminosity function in groups of a given mass is the same as that of the total galaxy population. After all, the conditional luminosity function of galaxies in groups varies significantly with group mass (Yang et al. 2005a; Zheng et al. 2005). Therefore, in the following iterations we use the group catalog of the previous iteration to self-calibrate $f(L_{19.5}, L_{lim})$. To do so, we first select all groups with $z < 0.09$ (for which $f = 1$) and compute their characteristic luminosities.

Next, we use these groups to determine the fraction of the characteristic luminosity that is contributed by group members with $L \geq L_{lim}$ for different values of L_{lim}. The top panels of Figure 4 show the results for three different values of L_{lim} (corresponding to $^{0.1}M_{r, lim} - 5 \log h = -20.0$, -20.5, and -21.0, from left to right). Next we determine the mean values of these fractions as function of $L_{19.5}$, which are shown in Figure 4 as open circles, and we define $f(L_{19.5}, L_{lim})$ as the exponential function that best fits these mean values (shown as solid lines in Fig. 4). Note, however, that the scatter around these mean values is fairly large. Consequently, despite the fact that our correction factors are self-calibrated in an iterative way, they are only accurate in a statistical
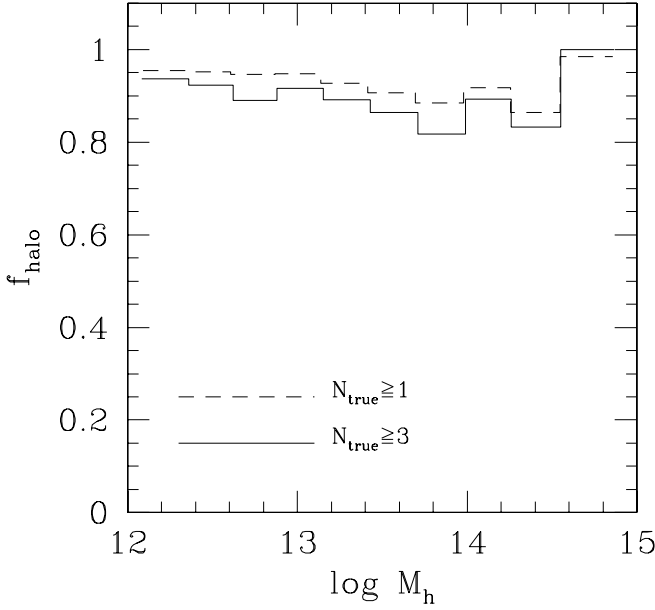
FIG. 3.—Global completeness, f_{halo}, defined as the fraction of halos in the MGRS whose brightest member has actually been identified as the brightest (central) galaxy of its group, as a function of the true halo mass M_h. Results are shown for all halos (*dashed histogram*) and for those halos with at least three members in the MGRS (*solid histogram*).

sense and are not expected to be accurate for individual groups. As we show in § 3.5, a considerable amount of scatter in the halo masses can be introduced by such correction.

Similar to the characteristic luminosity defined above, we also define a characteristic stellar mass

$$M_{\mathrm{stellar}} = \frac{1}{g(L_{19.5}, L_{\mathrm{lim}})} \sum_i \frac{M_{*,i}}{\mathcal{C}_i}, \qquad (13)$$

where as for the characteristic luminosity the summation is over all group members with $^{0.1}M_r - 5 \log h \leq -19.5$, and $g(L_{19.5}, L_{\mathrm{lim}})$ is a similar correction factor to $f(L_{19.5}, L_{\mathrm{lim}})$, but is tailored to the stellar mass rather than the r-band luminosity. Similar to f, these correction factors can be self-calibrated, and the results are shown in the bottom panels of Figure 4.

3.4. Correcting for Survey Edges

An additional incompleteness effect that needs to be accounted for is due to the survey geometry. A group whose projected area straddles one or more survey edges may have members that fall outside of the survey, thus causing an incompleteness, which in turn affects our mass estimate of the group. The geometry of the survey used here is defined as the region on the sky where the SDSS redshift completeness $\mathcal{C} > 0.7$. Clearly, this geometry is fairly complicated, which can potentially have a significant impact on various statistics of the group catalog (cf. Cooper et al. 2005;
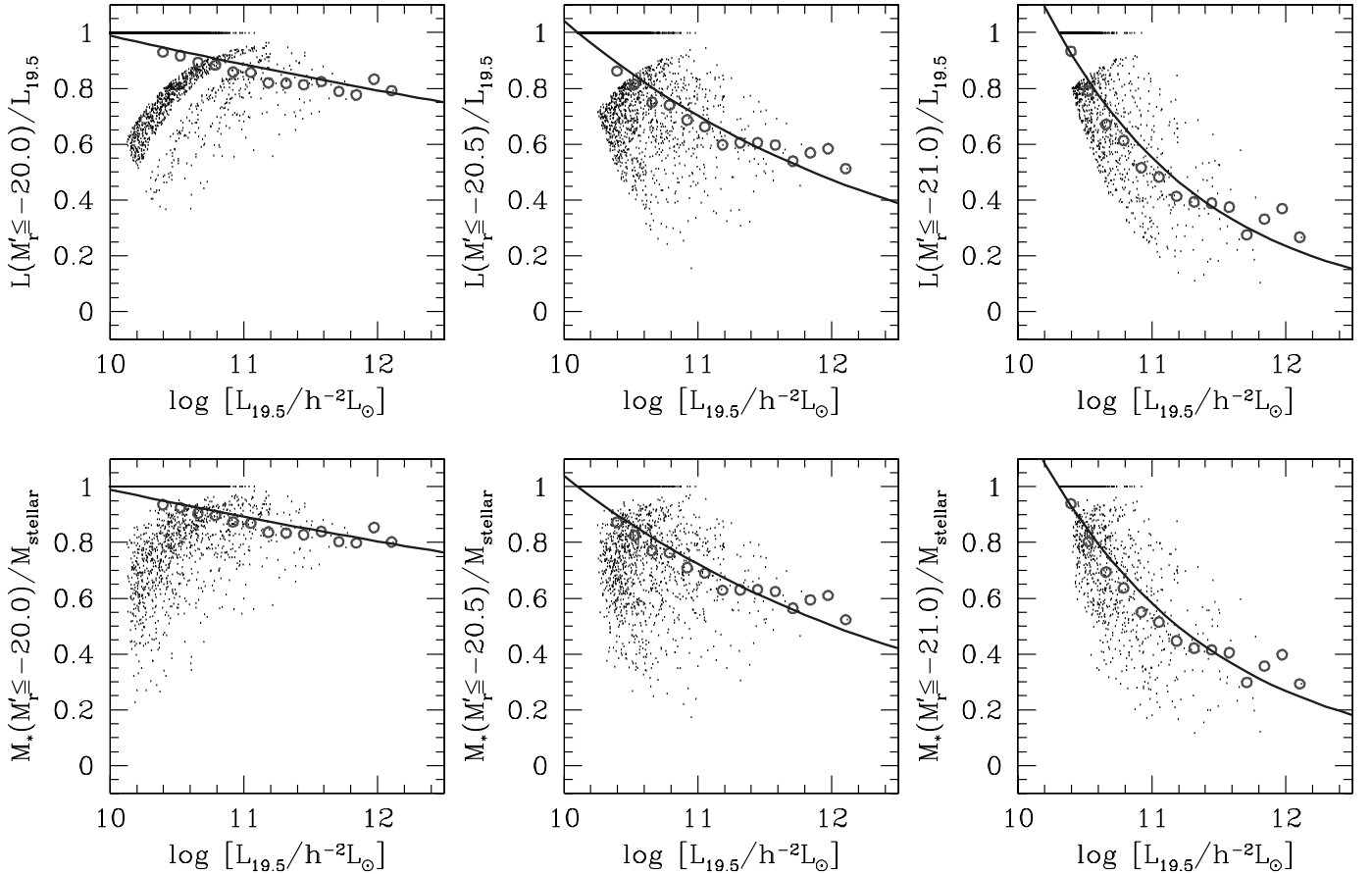

FIG. 4.—*Top*: Fraction of the characteristic luminosity $L_{19.5}$ that is contributed by galaxies above a given magnitude limit as a function of $L_{19.5}$. Left, middle, and right panels correspond to magnitude limits of $M_r' \equiv {}^{0.1}M_r - 5 \log h = -20.0$, -20.5, and -21.0, respectively. Each dot corresponds to a group in our SDSS group catalog based on sample II with $z < 0.09$. The open circles indicate the mean fractions for a given bin in $L_{19.5}$, while the solid line is the exponential function that best fits these mean values and which defines our completeness correction factor $f(L_{19.5}, L_{\mathrm{min}})$ discussed in the text. *Bottom*: Same as the top panels, except that here we plot the fraction of characteristic stellar mass, M_{stellar}, contributed by galaxies above a given magnitude limit. [*See the electronic edition of the Journal for a color version of this figure.*]
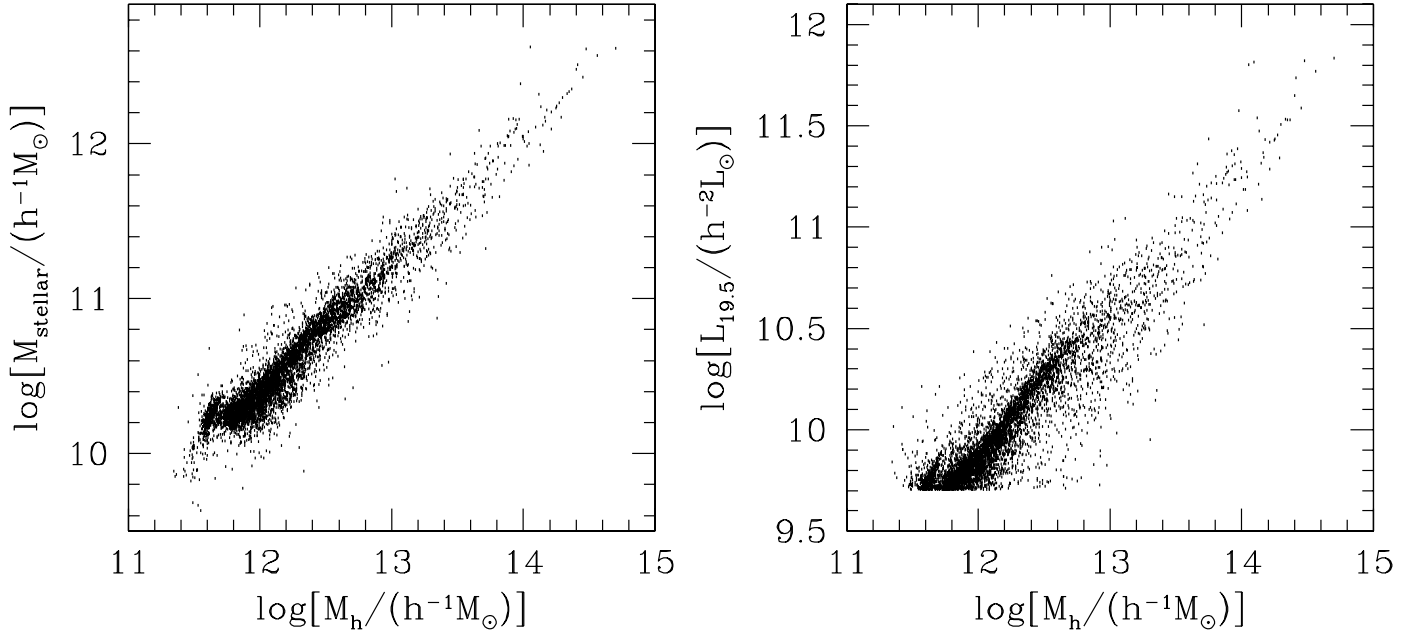
FIG. 5.—Distributions of characteristic stellar mass $M_{stellar}$ (total stellar mass of galaxies with $^{0.1}M_r - 5\log h \leq -19.5$ in a halo; *left*) and luminosity $L_{19.5}$ (total luminosity of galaxies with $^{0.1}M_r - 5\log h \leq -19.5$ in a halo; *right*) as functions of halo mass, M_h. The results are obtained from the semianalytical model of Kang et al. (2005). Obviously, both $M_{stellar}$ and $L_{19.5}$ are tightly correlated with the halo mass.

Berlind et al. 2006). In order to correct for these edge effects we proceed as follows.

First, we estimate the mass for each group using the method described in § 3.5 below without taking edge effects into account. We then randomly distribute 200 points within the corresponding halo radius r_{180} (which we compute using eq. [5]). Next we apply the SDSS DR4 survey mask and remove those random points that fall outside of the region where $\mathcal{C} > 0.7$. For each group we then compute the number of remaining points, N_{remain}, and we define $f_{edge} = N_{remain}/200$ as a measure for the volume of the group that lies within the survey edges. Finally, we multiply $L_{19.5}$ and $M_{stellar}$ with $1/f_{edge}$ to correct for the "missing members" outside of the edges of the survey. Tests with MGRSs show that this correction works well, except for groups with a small f_{edge}. We therefore discard those groups with $f_{edge} < 0.6$, which removes (only) 1.6% of all groups. After this correction for edge effect, we recalculate the mass for each group as described in § 3.5. The mass difference before and after this edge effect correction is relatively small: in most cases less than 3% and on average less than 10%. Since this change in group mass translates only in very small changes in r_{180}, no iteration of this procedure is required.

3.5. *Estimating Group Masses*

An important aspect of each galaxy group catalog is the determination of the masses of the groups. Most studies infer the (dynamical) group mass from the velocity dispersion of their member galaxies. However, the vast majority of the groups in our sample contain only a few members, making a dynamical mass estimate based on its members extremely unreliable. Mass estimates based on gravitational lensing (either strong or weak) or on X-ray emission also can only be applied to the most massive systems. Furthermore, these latter two methods require high-quality data in excess of the information directly available from the redshift survey used to construct the group catalog, rendering them impractical.

Rather, we estimate the group masses from their characteristic luminosities or characteristic stellar masses. This has the advantage that (1) it is equally applicable to groups spanning the entire range in richness, and (2) it does not require any additional data. As demonstrated in Yang et al. (2005b), the mass of a dark matter halo associated with a group is tightly correlated with the total luminosity of all member galaxies down to some luminosity. This is further illustrated in Figure 5, where we plot the correlations between the halo mass, M_h, and the characteristic stellar mass $M_{stellar}$ (*left*) and characteristic luminosity $L_{19.5}$ (*right*) in the semi-analytical model of Kang et al. (2005). Clearly, both $M_{stellar}$ and $L_{19.5}$ are tightly correlated with halo mass, with the $M_{stellar}$-M_h relation being slightly tighter than that between $L_{19.5}$ and M_h. This is expected, since $M_{stellar}$ is less affected by the current amount of star formation, and suggests that the characteristic stellar mass is a somewhat better mass indicator than the characteristic luminosity. On the other hand, the luminosities are directly observed, while the stellar masses are derived quantities, which creates additional scatter. Therefore, we compute two mass estimates for each group: M_S, based on the characteristic stellar mass $M_{stellar}$, and M_L, based on the characteristic luminosity $L_{19.5}$. Throughout, we compare all results from the group catalog for both mass estimates.

In order to convert the characteristic luminosities and stellar masses to halo masses, we make the assumption that there is a one-to-one relation between $L_{19.5}$ (or $M_{stellar}$) and M_h. For a given (comoving) volume and a given halo mass function, $n(M_h)$, one can then link the characteristic luminosity or stellar mass to a halo mass by matching their rank orders. Note, however, that this only works for a group sample that is complete in $L_{19.5}$ or $M_{stellar}$. In Figure 6 we plot the redshift distributions of groups in three different ranges of $L_{19.5}$ (*top*) and $M_{stellar}$ (*bottom*). Comparing these distributions with that expected for a constant number density (shown as the solid line), we obtain the rough redshifts out to which these different samples are complete. In Table 2 we list the redshift limits thus obtained for the three different bins of mass indicators, along with the numbers of groups in each of the complete samples. Only groups in these complete samples are used in the ranking; the masses of other groups are estimated by linear interpolation of the relations between M_h and each of the mass
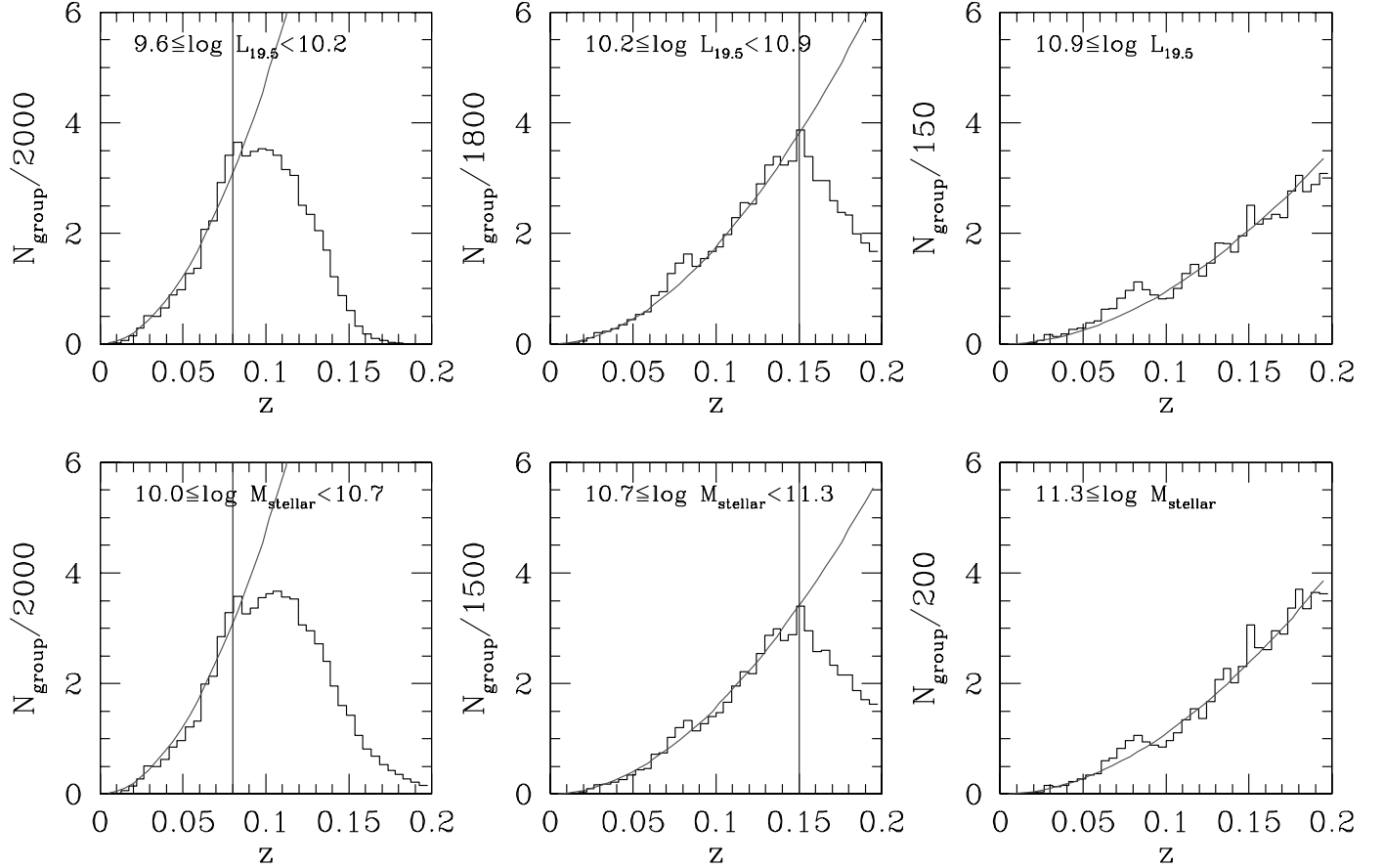
Fig. 6.—Redshift distributions of groups for three bins in characteristic luminosity $L_{19.5}$ (*top*) and characteristic stellar masses $M_{stellar}$ (*bottom*), as indicated. Solid lines indicate the expected values for a constant group number density. Whenever the observed distribution of groups starts to systematically drop below this line, we consider the sample incomplete. The vertical lines indicate the redshifts out to which we consider the samples complete. In the right panels, the group samples are considered complete out to the redshift limit of our galaxy sample ($z = 0.2$). [*See the electronic edition of the Journal for a color version of this figure.*]

indicators obtained from the complete samples. Because of the particular volume-limited samples used, we can assign group masses down to $10^{11.6}\ h^{-1}\ M_\odot$.

Clearly, the assumption of a one-to-one relation between the characteristic luminosity or stellar mass and the halo mass is over-simplified. In reality, these relations contain some scatter, which results in errors in our inferred group masses. However, detailed tests with mock galaxy redshift surveys have shown that this method nevertheless allows for a very accurate recovery of *average* halo occupation statistics. In particular, the group finder yields average halo occupation numbers and average mass-to-light ratios that are in excellent agreement with the input values (Yang et al. 2005c; Weinmann et al. 2006a). An additional shortcoming of our method is it requires the halo mass function, which is cosmology

dependent (e.g., Sheth et al. 2001; Warren et al. 2006).[8] However, as we show in § 4.1, it is extremely easy to convert the group masses to another cosmology, without having to rerun the group finder.

In order to further assess the reliability of the halo masses assigned to individual groups, we use the mock group catalog obtained from the CLF-based mock. Following the procedure described above we assign each (mock) group a halo mass M_L based on its ranking of the characteristic group luminosity $L_{19.5}$. The top left panel of Figure 7 shows the M_L thus obtained versus the true

---

[8] Throughout we compute the halo mass function using the formulae given in Warren et al. (2006) with the transfer function given by Eisenstein & Hu (1998), which properly accounts for the effects of baryons.

TABLE 2
COMPLETE SAMPLES USED FOR MASS RANKING

Redshift (1)	$\log L_{19.5}$ (2)	Groups Samples I/II/III (3)	$\log M_{stellar}$ (4)	Groups Samples I/II/III (5)
$0.01 \le z \le 0.20$	≥ 10.9	7583/7683/8409	≥ 11.3	11740/11851/12012
$0.01 \le z \le 0.15$	[10.2, 10.9]	75120/75306/66001	[10.7, 11.3]	53248/53377/47953
$0.01 \le z \le 0.08$	[9.6, 10.2]	33898/33939/36038	[10.0, 10.7]	32739/32789/32702

NOTES.—Properties of the three complete samples used to estimate group mass via the ranking of characteristic luminosity or characteristic stellar mass. Column (1) lists the redshift range of each sample. Columns (2) and (4) list the corresponding ranges in $\log L_{19.5}$ and $\log M_{stellar}$, respectively. Finally, columns (3) and (5) list the corresponding numbers of groups in the catalogs based on galaxy samples I, II, and III, respectively.
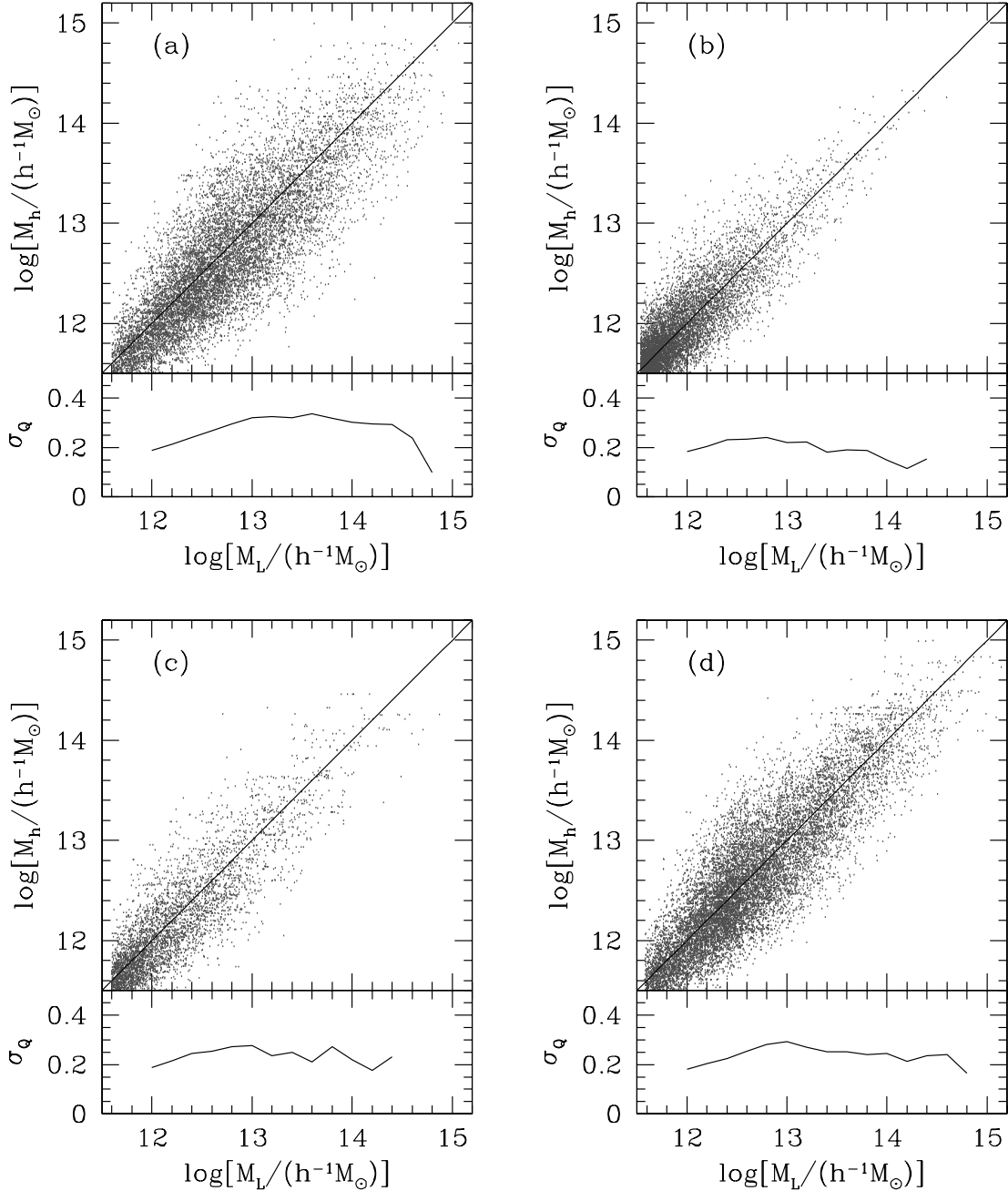
FIG. 7.—(*a*) Comparison between the assigned halo mass M_L, based on the characteristic group luminosity $L_{19.5}$, and the true halo mass M_h. These results are obtained from the mock group catalog constructed from our MGRS. The smaller bottom panel plots the standard deviation in Q defined by eq. (14) and reflects the amount of scatter with respect to the line of equality $M_L = M_h$ (shown as a solid line in the scatter plot). (*b*) Same as (*a*), except that here we show the comparison between M_h and the assigned halo mass M_L estimated from the ranking of the characteristic group luminosity $L_{19.5}$ obtained *directly* from the true group members in the simulation box used to construct the MGRS. (*c*) Same as (*a*), but this time only plotting the results for groups with $z \leq 0.09$ for which one does not need to correct $L_{19.5}$ for missing members. (*d*) Same as (*a*), but where we have mimicked a perfect group finder without interlopers ($f_i = 0$) and with a completeness $f_c = 1$. See text for a detailed discussion. [*See the electronic edition of the Journal for a color version of this figure.*]

halo mass, M_h, defined as the mass of the dark matter halo that hosts the brightest group galaxy. In order to quantify the scatter with respect to the line of equality ($M_L = M_h$), we determine for each group the quantity

$$Q \equiv \frac{1}{\sqrt{2}} \left[\log (M_L) - \log (M_h) \right] \qquad (14)$$

and measure the standard deviation, σ_Q, in several bins of $[\log (M_L) + \log (M_h)]/2$. The results, shown in the bottom of the panel, indicate that the scatter is ∼0.35 dex for groups with

10^{13} h^{-1} $M_\odot \lesssim M_L \lesssim 10^{14.5}$ h^{-1} $M_\odot$, dropping to ∼0.2 dex at the high- and low-mass ends.

There are several factors that contribute to this scatter. The first is the intrinsic scatter in the relation between the halo mass and the true value of $L_{19.5}$. The top right panel of Figure 7 shows the relation between the true halo mass and the assigned mass based on the ranking of the true $L_{19.5}$ obtained from the CLF mock before incorporating any observational effects (e.g., magnitude limit, incompleteness, and survey boundary). In other words, we measure $L_{19.5}$ using all mock galaxies with $^{0.1}M_r - 5 \log h \leq -19.5$, independent of whether those galaxies would be incorporated in the

mock survey or not. The resulting scatter is about 0.2 dex, similar to that of the semianalytical model shown in the right panel of Figure 5.

The second source of scatter owes to the incompleteness and contamination introduced by our group finder. The bottom left panel of Figure 7 shows the relation between the assigned mass and the true mass for groups with $z < 0.09$. As discussed in § 3.3, the characteristic luminosity of these groups does not need to be corrected for incompleteness due to the magnitude limit of the survey (i.e., all galaxies with $^{0.1}M_r - 5 \log h \leq -19.5$ make the magnitude limit of the survey). The scatter here is only marginally larger than the intrinsic scatter shown in the top right panel, suggesting that the group-finding algorithm by itself only introduces a very small amount of uncertainty in the assigned masses.

The final source of scatter in the assigned group masses owes to the fact that for groups with $z > 0.09$ we need to correct the characteristic luminosity for the group members that do not meet the magnitude limit of the survey. As shown in Figure 4, this can introduce a considerable amount of scatter. To assess its impact on the inferred halo masses we proceed as follows. We group all galaxies in the mock SDSS DR4 according to the halo to which they belong. This resulting "group catalog" has, by construction, a completeness $f_c = 1$, an interloper fraction $f_i = 0$, and a purity $f_p = 1$. For each group in this perfect catalog we estimate the characteristic luminosity $L_{19.5}$: for groups with $z < 0.09$ we simply sum the luminosities of all galaxies with $^{0.1}M_r - 5 \log h \leq -19.5$, while for groups with $z > 0.09$ we use the correction factors $f(L_{19.5}, L_{\lim})$ as described in § 3.3. Finally, we assign each group a mass M_L based on the ranking of $L_{19.5}$ as described above. The bottom right panel of Figure 7 plots the resulting M_L as a function of the true halo mass M_h. The scatter is $\sim$0.25 dex, comparable to that in the bottom left panel.

We therefore conclude that the majority of the scatter in the relation between the true and assigned halo masses owes to the intrinsic scatter in the relation between halo mass and its characteristic luminosity. The fact that the group finder is not perfect (i.e., suffers from interlopers and incompleteness) and that we need to correct the characteristic luminosity for members that do not make the magnitude limit of the survey only adds a relatively small contribution to the total scatter.

4. BASIC PROPERTIES OF THE GROUP CATALOG

Application of our halo-based group finder to the SDSS DR4 data set described in § 2 results in 295,992, 301,237, and 300,049 groups for samples I, II, and III, respectively. In what follows we present a few global properties of these group catalogs.

Table 3 lists, for each of the three samples, the number of groups with 1, 2, 3, and more than 3 members; clearly, the majority of the groups contain only a single member. Note also that sample III yields many more systems with richness $N \geq 2$ than samples I and II; this is simply due to the fact that almost all 38,672 galaxies with an assigned redshift are members of such systems. As shown in Zehavi et al. (2002), about 40% of these assigned redshifts have an error of more than 500 km s^{-1}. This means that in most cases these galaxies should not have been assigned to the group in question (i.e., they are interlopers), which obviously causes a systematic bias toward too many members per group. On the other hand, not taking account of the galaxies lost because of fiber collisions results in an opposite bias toward too few members per group. We can assess the impact of these biases on the group catalog by comparing results obtained from samples II and III. We come back to this issue below in this section.

As an illustration, Figure 8 shows the distributions of galaxies and groups in a 3° slice. As expected, massive groups are located

TABLE 3
Number of Galaxies and Groups in the SDSS DR4

Samples (1)	Galaxies (2)	Groups (3)	$N = 1$ (4)	$N = 2$ (5)	$N = 3$ (6)	$N > 3$ (7)
I......................	362,356	295,992	266,763	19,522	4511	5196
II....................	369,447	301,237	271,420	19,868	4619	5330
III...................	408,119	300,049	250,492	33,537	7848	8172

Notes.—For each of the three samples, columns (2) and (3) list the number of galaxies and of groups, respectively. In addition, columns (4)–(7) list the numbers of groups with 1, 2, 3, and more than 3 members.

in the denser regions of the galaxy density field, while groups with lower masses are more diffusely distributed. The clustering properties of these groups directly reflect the clustering properties of dark matter halos and can thus be used to directly probe the mass dependence of the halo bias (cf. Yang et al. 2005c; Coil et al. 2006; Berlind et al. 2006). We defer a more detailed analysis of the clustering properties of the groups in the SDSS DR4 group catalogs presented here to a forthcoming paper.

Figure 9 plots the number of groups as a function of group richness (*left*), redshift (*middle*), and halo mass (*right*). Group redshift is estimated using the luminosity-weighted average of all member galaxies. Dashed, solid, and dotted histograms correspond to the group catalogs based on samples I, II, and III, respectively. As already mentioned above, groups obtained from sample III are systematically richer than those obtained from the other two samples, which simply owes to the fact that all galaxies with an assigned redshift are group members. However, as is evident from the middle and right panels of Figure 9, the redshift distributions and halo mass functions of all three group samples are extremely similar; although the inclusion of galaxies with assigned redshifts changes the richness of the systems, their redshifts and inferred masses are virtually unaffected. The long-dashed line in the right panel of Figure 9 shows the theoretical mass function of dark matter halos over the redshift range $0.01 \leq z \leq 0.20$. The mismatch between the group mass function and this theoretical halo mass function at $\log(M_L/h^{-1} \ M_\odot) \lesssim 10^{12.8}$ is caused by the incompleteness of the group catalog shown in Figure 6 and discussed in § 3.5. If we would only plot the mass functions for groups in the complete samples of Table 2, they would, by construction, perfectly match their theoretical equivalent.

As discussed in § 3.5, group masses are obtained down to $10^{11.8} \ h^{-1} \ M_\odot$ using two different mass indicators, the characteristic luminosity $L_{19.5}$ and the characteristic stellar mass $M_{\rm stellar}$. The left panel of Figure 10 compares the inferred group masses M_L and M_S, obtained using $L_{19.5}$ and $M_{\rm stellar}$, respectively. Overall, both halo masses agree very well with each other, with an average scatter that decreases from $\sim$0.1 dex at the low-mass end to $\sim$0.05 dex at the massive end. This scatter is expected and mainly reflects that galaxies of a given luminosity have different colors and therefore different (inferred) stellar masses. The effect is somewhat larger for lower mass groups, simply because their characteristic mass and luminosity are dominated by a smaller number of galaxies.

Finally, to assess how the uncertainties in the correction for fiber collisions affect the group catalog, we compare the masses of groups in sample II with those of its counterparts in sample III. Here a group in sample III is defined as the counterpart of a group in sample II if it has the same brightest (central) galaxy. We can find a sample III counterpart for $\sim$95% of all groups in sample II. There are two main reasons why a group may not have a counterpart. First of all, in about 3% of the groups in sample III, the
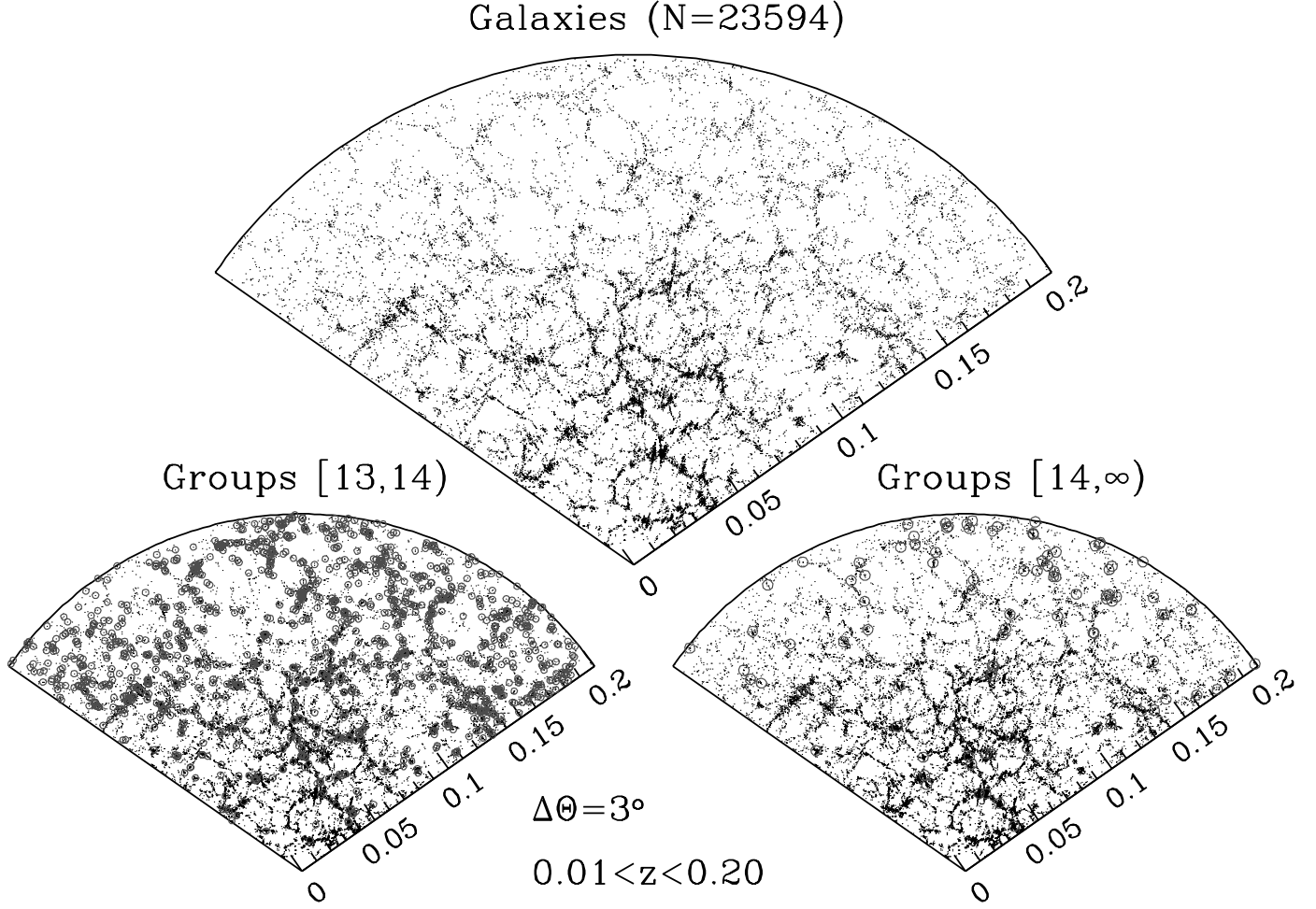
FIG. 8.—Distribution of a subset of SDSS DR4 galaxies in a 3° slice in the south galactic pole region of the SDSS is shown in the large wedge. These distributions are repeated in the smaller wedges, where we overplot, as open circles, the groups with assigned masses in the range $10^{13}\,h^{-1}\,M_\odot$ to $10^{14}\,h^{-1}\,M_\odot$ (*bottom left wedge*) and $>10^{14}\,h^{-1}\,M_\odot$ (*bottom right wedge*). Note the halo masses used in this plot are obtained from the ranking of the characteristic luminosity $L_{19.5}$. [*See the electronic edition of the Journal for a color version of this figure.*]
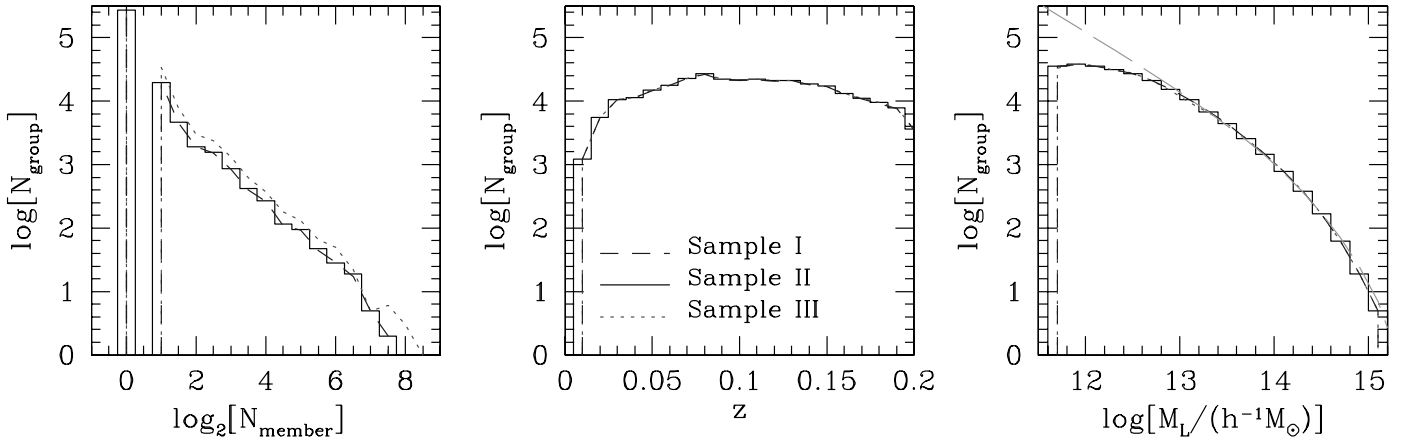
FIG. 9.—Number of groups as a function of the number of group members (*left*), group redshift (*middle*), and assigned halo mass (*right*). The dashed lines, histograms, and dotted lines show the results for the group catalogs based on samples I, II, and III, respectively. For comparison, in the right panel we also plot the theoretical halo mass function (*long-dashed line*). For $\log(M_L/h^{-1}\,M_\odot) \gtrsim 13$ the mass function of the groups is in excellent agreement with this theoretical mass function, indicating that our group sample is complete for this mass range. For lower mass groups, the sample is only complete out to lower redshifts (cf. Table 2 and Fig. 6). Note that in the right panel the group masses have been assigned based on their characteristic luminosity $L_{19.5}$. Using the characteristic stellar mass, $M_{\rm stellar}$, instead results in an almost identical plot. [*See the electronic edition of the Journal for a color version of this figure.*]
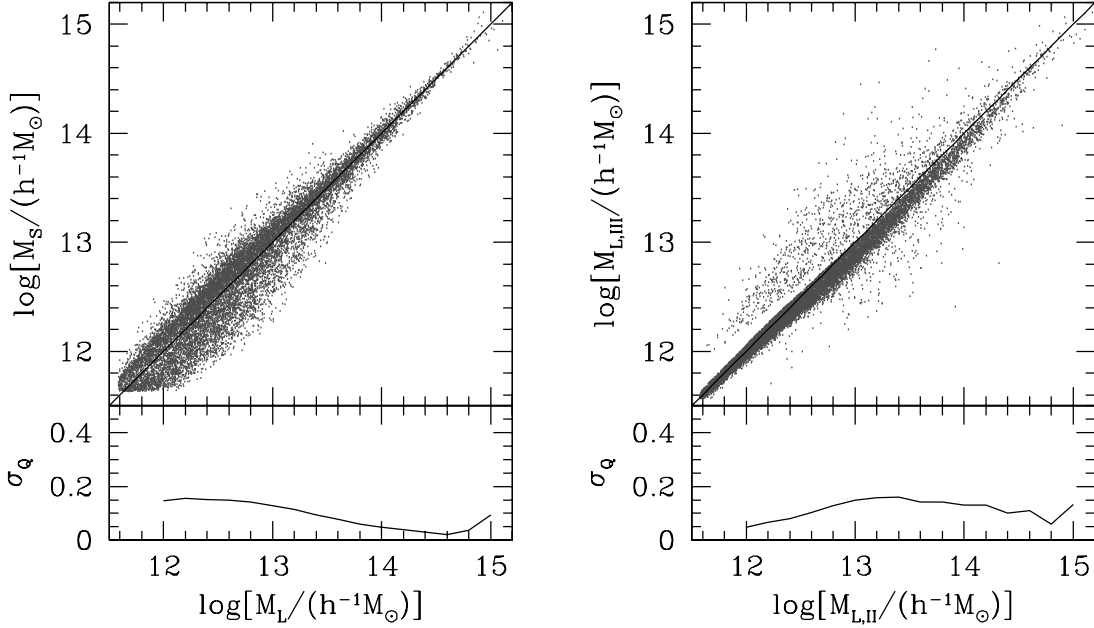
FIG. 10.—*Left*: Comparison of the group masses M_L and M_S, obtained using the two different mass indicators, $L_{19.5}$ and M_{stellar}, respectively. As expected, both masses agree extremely well: the standard deviation in Q, shown in the bottom panel and defined by eq. (14), is less than 0.05 dex at the massive end. At the low-mass end, $\sigma_Q \sim 0.1$ dex due to the smaller average number of galaxies per group. *Right*: Halo mass assigned to a group in sample II vs. the halo mass of the corresponding group in sample III. Here a group in sample III is defined as the counterpart of a group in sample II if and only if it has the same brightest galaxy. Roughly 95% of all groups in sample II with $M_L \gtrsim 10^{12}\ h^{-1}\ M_\odot$ have a counterpart in sample III, and for more than 90% of these systems, the difference in the assigned group mass is smaller than 50%. There is a small number of outliers, but they do not have a significant effect on the overall statistical properties of the group samples. [*See the electronic edition of the Journal for a color version of this figure.*]

brightest group member is actually a galaxy with an assigned redshift, so that this group cannot have a counterpart in sample II. In addition, about 1% of the groups in sample II merge with other (nearby) groups when the additional galaxies with assigned redshifts are used. Consequently, in terms of their brightest galaxies, some groups in sample II have disappeared in sample III (i.e., do not have a counterpart), while others have suddenly increased their mass substantially because they have now merged with another group. The right panel of Figure 10 plots the relation between the assigned halo mass of a group in sample II and that of its counterpart in sample III. The relation is extremely tight: for more than 90% of the systems, the difference in the assigned group mass is smaller than 50%, at any given mass. This emphasizes once again that although the incompleteness due to fiber collisions can have a significant effect on the richness of individual groups, it does not have a significant impact on which systems are selected as groups or on their assigned masses. Although the standard deviation in Q (defined by eq. [14] but with M_L and M_h replaced by $M_{L,\text{II}}$ and $M_{L,\text{III}}$, respectively) reaches up to $\sim$0.25 dex, this largely owes to a very small fraction of outliers that are clearly visible in the scatter plot. In particular, one can distinguish a second "sequence" of systems with $M_{L,\text{III}} > M_{L,\text{II}}$; this corresponds to the 1% of groups in sample II mentioned above that have merged due to the additional galaxies with assigned redshifts.

4.1. *Average Mass-to-Light Ratios*

The mass-to-light ratio of a dark matter halo expresses the efficiency with which stars have formed in that halo. Consequently, accurate measurements of the average mass-to-light ratios of dark matter halos as a function of halo mass can put tight constraints on the physics of galaxy formation. The top left panel of Figure 11 shows the average mass-to-light ratios, $M_h/L_{19.5}$, as a function of halo mass obtained from our group catalog. Results are shown based on both mass indicators described above, $L_{19.5}$ (*solid line*) and M_{stellar} (*open circles*). The error bars indicate the 68th percen-

tiles from the distributions in each M_S mass bin. Since M_L is based on the ranking of $L_{19.5}$, there is no scatter in the corresponding mass-to-light ratios. As one can see, the mass-to-light ratios obtained from the two different mass indicators are in extremely good agreement with each other. Note that the results shown here are obtained from sample II; those for samples I and III are again very similar and, consequently, are not shown for the sake of argument. For reference, all mass-to-light ratios obtained from sample II are listed in Table 4. The mass-to-light ratios have also been obtained by various observations, e.g., Carlberg et al. (1996) from the CNOC sample and Popesso et al. (2004) from RASS-SDSS. We defer a more detailed comparison to previous results in a forthcoming paper.

In addition to the mass-to-light ratios, $M_h/L_{19.5}$, we can also use our group catalogs to compute the ratio between halo mass and characteristic stellar mass, M_h/M_{stellar}. The bottom left panel of Figure 11 shows the average halo mass–to–stellar mass ratios, M_h/M_{stellar}, as a function of halo mass. Once again we show the results obtained from both mass indicators. This time the open circles with error bars reflect the results obtained using $L_{19.5}$ as the mass indicator, while the results based on the characteristic stellar mass are shown as a solid line. Since M_S is based on the ranking of M_{stellar}, this time there is no scatter in the results based on the characteristic stellar mass, while the error bars reflect the 68th percentiles of the distributions in M_h/M_{stellar}, where M_h is obtained from $L_{19.5}$. As for the mass-to-light ratios, the results based on both mass indicators are extremely similar and are listed in Table 4.

Since our mass assignments require use of the halo mass function, which is cosmology dependent, it is important to investigate how the average $M_h/L_{19.5}$ and M_h/M_{stellar} change if we change cosmology. The dashed lines in the right panels of Figure 11 show the results obtained when adopting a WMAP1 cosmology ($\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$, $\sigma_8 = 0.9$) instead of a WMAP3 cosmology. Changing the cosmology changes (1) the luminosity and angular distances of all galaxies in the SDSS DR4, and thus their absolute
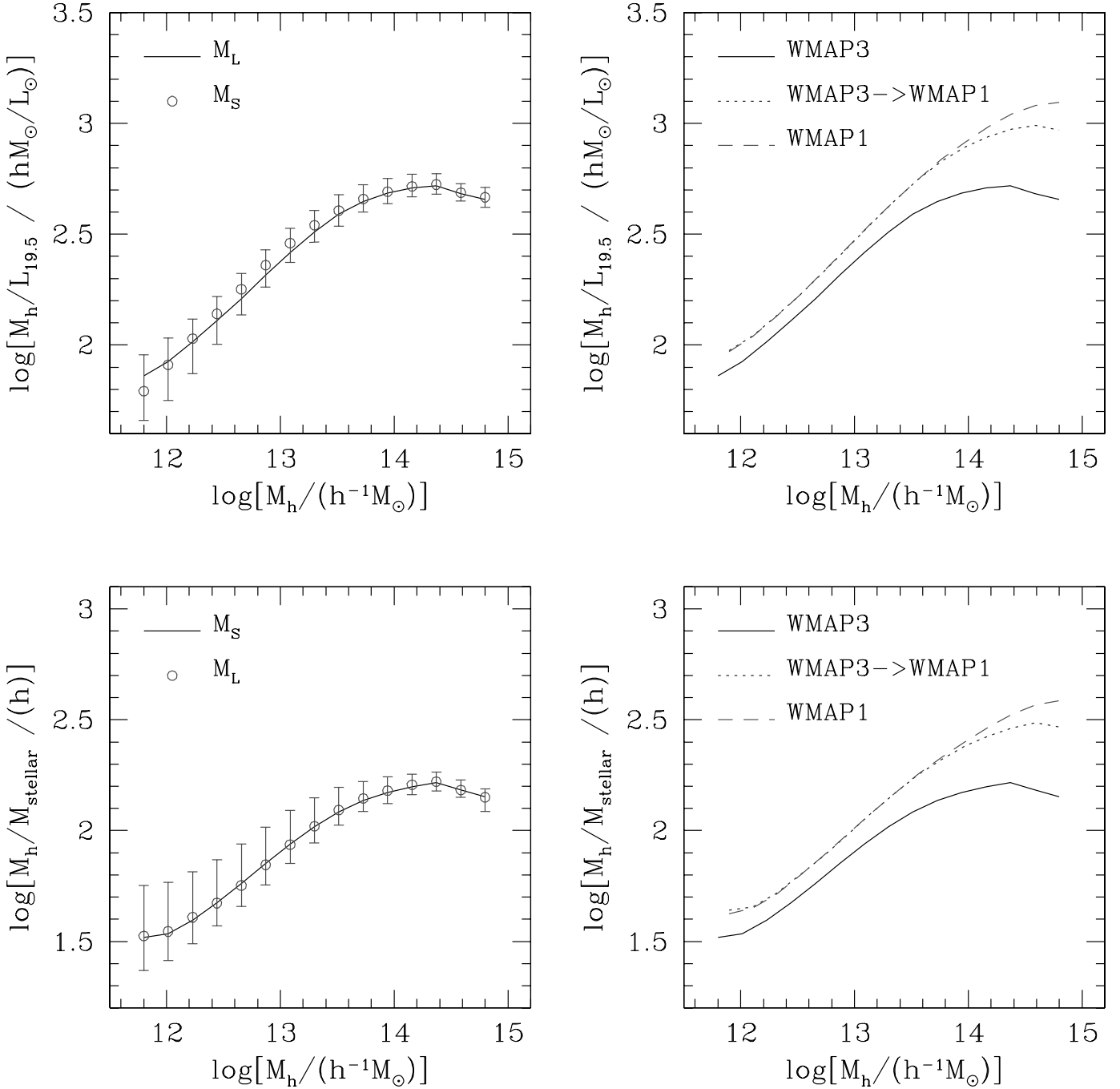
Fig. 11.—*Top left*: Inferred mass-to-light ratios, $M_h/L_{19.5}$, of galaxy groups in the SDSS DR4 as a function of their assigned halo mass, M_h (see also Table 4). The solid line and open circles correspond to the mass-to-light ratios obtained using the group masses M_L (based on the characteristic luminosity) and M_S (based on the characteristic stellar mass), respectively. By construction, there is no scatter in the inferred mass-to-light ratios when using M_L, since the characteristic luminosity is assumed to be related to the halo mass on a one-to-one basis. *Bottom left*: Same as the top left panel, except that this time we plot the inferred ratios between halo mass and characteristic stellar mass, $M_h/M_{stellar}$. This time there is no scatter when using M_S to infer the halo mass (*solid line*), and the error bars reflect the 68% confidence levels of $M_h/M_{stellar}$ obtained using M_L as mass indicator. *Right*: Solid lines are the same as in the corresponding left panels and are obtained assuming a WMAP3 cosmology throughout. The dashed lines show the results obtained when using a WMAP1 cosmology instead. Finally, the dotted lines correspond to the results obtained when using the group finder assuming a WMAP3 cosmology, but using a WMAP1 halo mass function when inferring the final group masses. See text for a detailed discussion. [*See the electronic edition of the Journal for a color version of this figure.*]

magnitudes and (comoving) separations, and (2) the halo mass function. The former has an almost negligible (small at massive end) effect, mainly because our sample of galaxies is restricted to $z < 0.2$. The halo mass function, however, has a strong impact. Since there are more massive halos in a WMAP1 cosmology than in a WMAP3 cosmology, the ranking assigns a larger halo mass to a given group, which results in larger values of $M_h/L_{19.5}$ and $M_h/M_{stellar}$.

Note that the mass-to-light ratios are also used in the group finder to assign memberships to groups (see § 3.1). This suggests that whenever one decides to change one or more cosmological parameters, one has to rerun the entire group finder in order to obtain new mass estimates (as we did for the WMAP1 and WMAP3 cosmologies shown in Fig. 11). This is impractical if one intends to use the group catalog to constrain different cosmological models. In order to test whether we can avoid having to rerun the group finder

TABLE 4
RATIOS BETWEEN HALO MASS AND LUMINOSITY AND BETWEEN HALO MASS AND STELLAR MASS

$\log(M_h/h^{-1}\,M_\odot)$ (1)	$\log(M_L/L_{19.5})$ (2)	$\log(M_S/L_{19.5})$			$\log(M_S/M_{\rm stellar})$ (6)	$\log(M_L/M_{\rm stellar})$		
		16% (3)	50% (4)	84% (5)		16% (7)	50% (8)	84% (9)
11.80..........................	1.860	1.659	1.793	1.957	1.518	1.335	1.491	1.741
12.00..........................	1.921	1.736	1.900	2.025	1.532	1.411	1.541	1.763
12.20..........................	2.003	1.777	1.988	2.099	1.587	1.479	1.600	1.807
12.40..........................	2.089	1.876	2.100	2.192	1.660	1.554	1.660	1.858
12.60..........................	2.183	2.043	2.213	2.291	1.738	1.633	1.731	1.921
12.80..........................	2.281	2.215	2.324	2.394	1.822	1.723	1.813	1.992
13.00..........................	2.379	2.332	2.421	2.491	1.906	1.805	1.895	2.058
13.20..........................	2.470	2.421	2.504	2.570	1.983	1.887	1.973	2.117
13.40..........................	2.551	2.500	2.573	2.641	2.052	1.964	2.049	2.165
13.60..........................	2.616	2.562	2.629	2.694	2.106	2.052	2.115	2.205
13.80..........................	2.662	2.615	2.672	2.737	2.151	2.100	2.160	2.231
14.00..........................	2.693	2.646	2.699	2.755	2.180	2.134	2.187	2.247
14.20..........................	2.714	2.675	2.718	2.775	2.203	2.170	2.211	2.259
14.40..........................	2.718	2.681	2.723	2.767	2.216	2.178	2.220	2.261
14.60..........................	2.680	2.645	2.681	2.718	2.181	2.145	2.179	2.216
14.80..........................	2.657	2.621	2.667	2.712	2.152	2.086	2.151	2.189

NOTES.—Column (1): logarithm of the assigned halo mass. Column (2): average of the logarithm of the ratio between the assigned halo mass M_L and the characteristic luminosity $L_{19.5}$. Columns (3)–(5): 16th, 50th, and 84th percentiles of the distributions of the logarithm of the ratio between the assigned halo mass M_S and the characteristic luminosity. Column (6): logarithm of the ratio between assigned halo mass M_S and the characteristic stellar mass $M_{\rm stellar}$. Columns (7)–(9): 16th, 50th, and 84th percentiles of the distributions of the logarithm of the ratio between the assigned halo mass M_L and the characteristic stellar mass. The mass-to-light ratios $M_L/L_{19.5}$ and $M_S/L_{19.5}$ are in units of $h\,M_\odot/L_\odot$, and the ratios between halo mass and the characteristic stellar mass $M_L/M_{\rm stellar}$ and $M_S/M_{\rm stellar}$ are in units of h. All these results correspond to a WMAP3 cosmology, and we emphasize once more that all luminosities are in the SDSS r-band and have been $K + E$ corrected to $z = 0.1$.

when changing cosmology we proceed as follows. We run our group finder over the SDSS DR4 assuming a WMAP3 cosmology, but then, when we convert $L_{19.5}$ or $M_{\rm stellar}$ into halo mass we use the WMAP1 halo mass function. The results are shown in the right panels of Figure 11 as dotted lines. Clearly, the impact of assuming a different cosmology in the group finder is almost negligible. This demonstrates that one can simply convert the $M_h/L_{19.5}$ and $M_h/M_{\rm stellar}$ listed in Table 4 to another cosmology (as long as it is not too different from the WMAP3 cosmology adopted here), without having to rerun the group finder over the data, by using the relation

$$\int_{M_h}^{\infty} n\left(M'_h\right) dM'_h = \int_{\widetilde{M}_h}^{\infty} \tilde{n}\left(M'_h\right) dM'_h, \qquad (15)$$

where M_h and $n(M_h)$ are the mass and halo mass function in the WMAP3 cosmology, respectively, and $\widetilde{M}_h$ and $\tilde{n}(M_h)$ are the corresponding values in the other cosmology.

4.2. Group Velocity Dispersions

For relatively massive groups, especially for groups with a sufficient number of member galaxies, one can estimate a dynamical group mass based on the velocity dispersion of the member galaxies. Following Yang et al. (2005b), we use the gapper estimator described by Beers et al. (1990) to estimate the line-of-sight velocity dispersion of each individual group. The method involves ordering the set of recession velocities $\{v_i\}$ of the N group members and defining gaps as

$$g_i = v_{i+1} - v_i, \quad i = 1, 2, \ldots, N - 1. \qquad (16)$$

The rest-frame velocity dispersion is then estimated by

$$\sigma_{\rm gap} = \frac{\sqrt{\pi}}{(1 + z_{\rm group})N(N - 1)} \sum_{i=1}^{N-1} w_i g_i, \qquad (17)$$

where the weight is defined as $w_i = i(N - i)$. Since there is a central galaxy in each group, which is assumed to be at rest with respect to the dark matter halo, the estimated velocity dispersion has to be corrected. This results in a final velocity dispersion given by

$$\sigma = \sqrt{\frac{N}{N - 1}} \sigma_{\rm gap}. \qquad (18)$$

The top panels of Figure 12 show the line-of-sight velocity dispersions of groups thus obtained as a function of the assigned halo mass M_L for groups with at least three (top left) and with at least eight members (top right). Solid triangles with error bars indicate the mean and the 1 σ scatter of the line-of-sight velocity dispersion in each mass bin. Clearly, there is a good correlation between the velocity dispersion and the mass M_L, indicating that the assigned masses are reliable mass indicators. Compared to the theoretical prediction of equation (6), which is shown as a solid line, the line-of-sight velocity dispersions of the group members are on average $\sim$40% lower. As discussed in Yang et al. (2005b), this discrepancy is mainly due to the fact that galaxies with the highest peculiar velocities in a group are the most likely to be missed by the group finder. To demonstrate this, the bottom panels of Figure 12 show the corresponding results obtained from our mock group catalog. In this case, the input velocity dispersions for galaxies in halos have a mean relation that is given by the solid line, and the halo masses M_h are the true halo masses. Here again, we see that the velocity dispersions among the selected group members are lower than those implied by the halo masses.

5. SUMMARY

In this paper we have used a modified version of the halo-based group-finding algorithm developed in Yang et al. (2005b) to construct group catalogs from the SDSS DR4. Changes and
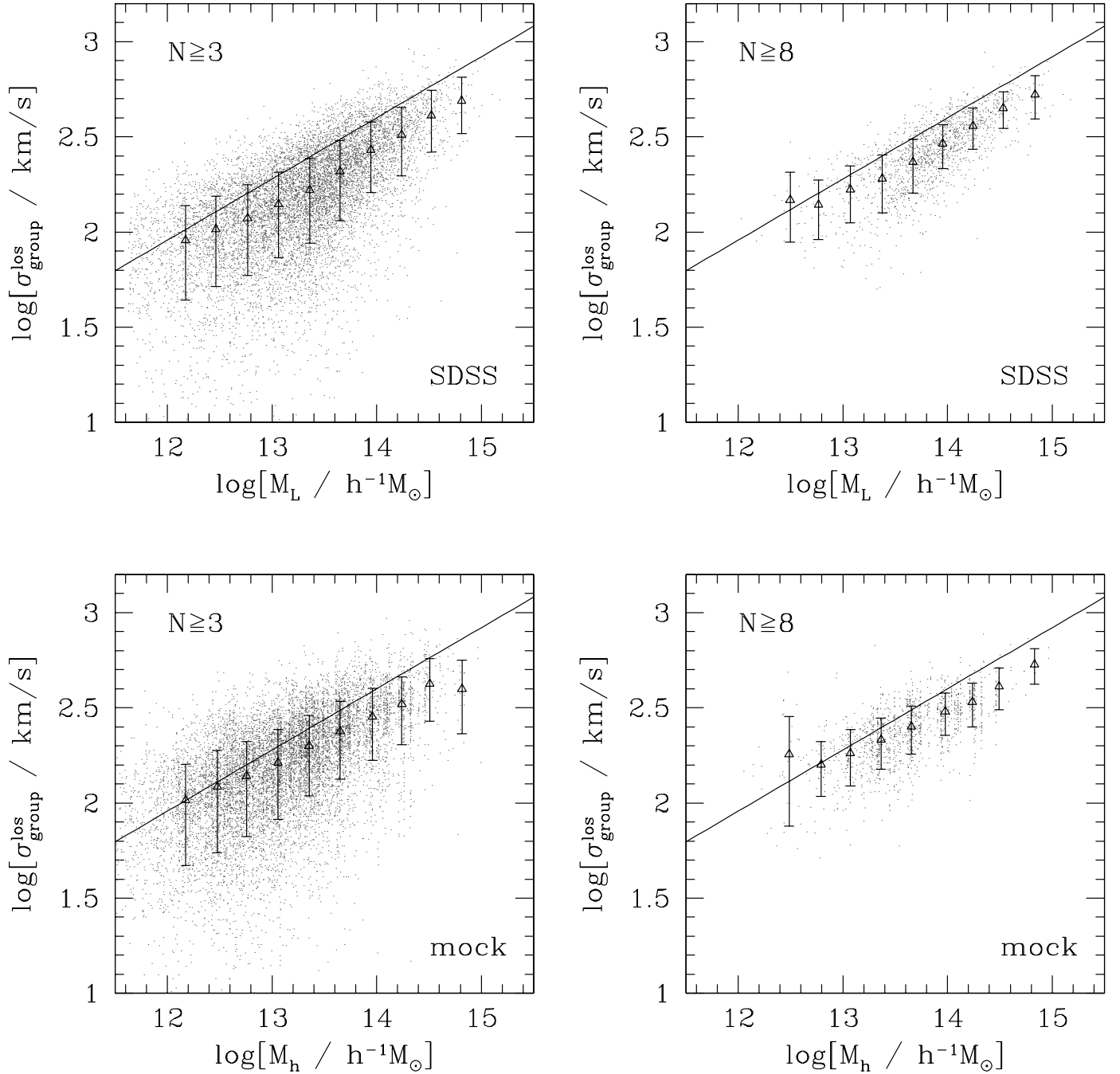
F_{IG}. 12.—Line-of-sight velocity dispersions of galaxies in groups, obtained using the gapper estimator of Beers et al. (1990), as function of halo mass. In the top panels we show results for groups in our SDSS DR4 group catalog as function of the assigned halo mass M_L. In the bottom panels the results have been obtained from the group catalog extracted from our MGRS and are shown as function of the true halo mass M_h. Left- and right-hand panels show the results for groups with at least three and at least eight members, respectively. Solid triangles with error bars indicate the mean and the 1 σ scatter in each mass bin, while the solid line reflects the theoretical expectation values based on eq. (6). As discussed in the text, the line-of-sight velocity dispersions of group members are biased low, due to the fact that galaxies with the highest peculiar velocities in a group are the most likely to be missed by the group finder. [*See the electronic edition of the Journal for a color version of this figure.*]

improvements in the group finder have been made in the following aspects:

1. In order to assign group memberships, we need to estimate masses for all tentative groups. Rather than using a model for the mass-to-light ratios, as we did previously, we now use self-consistent mass-to-light ratios obtained from the group catalog in an iterative way.

2. In order to correct the characteristic luminosity and stellar mass for missing members due to the magnitude limit of the sur-

vey, we use the mean correction factors obtained self-consistently from groups at low redshifts.

3. We have corrected the survey edge effect on the groups by a correction factor.

4. In order to estimate group masses, we use two different mass indicators, one based on the characteristic luminosity, $L_{19.5}$, and the other on the characteristic stellar mass, $M_{\rm stellar}$.

Tests based on detailed mock SDSS DR4 catalogs show that ~80% of all groups have a completeness >80%. The fraction of groups with a completeness of 100% ranges from ~60%

for the most massive groups to >95% for groups with masses in the range $10^{12.5}$ h^{-1} $M_\odot < M_h < 10^{13}$ h^{-1} $M_\odot$. On the order of 85% of all groups have an interloper fraction <50%, while ~65% of the groups have zero interlopers.

We have applied our group finder to three galaxy samples constructed from the SDSS DR4 galaxy catalog: sample I, which only contains galaxies with measured redshifts from the SDSS; sample II, which also contains those SDSS galaxies for which redshifts are available from alternative sources (mainly from the 2dFGRS); and sample III, which also includes galaxies which due to fiber collisions do not have a measured redshift, but which have been assigned the redshift of their nearest neighbor. We obtain a total of 295,992, 301,237, and 300,049 groups from samples I, II, and III, respectively, and each group is assigned two values for its halo mass based on the ranking of either the characteristic luminosity or the characteristic stellar mass of its member galaxies.

In this paper we have presented some of the basic properties of the group catalog, such as the distributions of richness, redshift, and mass. In addition, we have presented the average ratios between halo mass and characteristic luminosity and between halo mass and characteristic stellar mass. Although these are cosmology dependent, we have demonstrated that it is straightforward to convert these to other cosmologies. A more detailed analysis of the group properties and their implications for halo occupation statistics, galaxy formation, and cosmology will be presented in a series of forthcoming papers. As a final note, we mention that all group catalogs presented here are available from the authors on request.

REFERENCES

Abazajian, K., et al. 2004, AJ, 128, 502
———. 2005, AJ, 129, 1755
Adelman-McCarthy, J. K., et al. 2006, ApJS, 162, 38
Annis, J., et al. 1999, BAAS Abstr., 195, 1202
Bahcall, N. A., et al. 2003, ApJS, 148, 243
Baldry, I. K., Glazebrook, K., Brinkmann, J., Ivezic, Z., Lupton, R. H., Nichol, R. C., & Szalay, A. S. 2004, ApJ, 600, 681
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289
Berlind, A. A., Kazin, E., Blanton, M. R., Pueblas, S., Scoccimarro, R., & Hogg, D. H. 2006, ApJ, submitted (astro-ph/0610524)
Berlind, A. A., & Weinberg, D. H. 2002, ApJ, 575, 587
Berlind, A. A., et al. 2006, ApJS, 167, 1
Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734
Blanton, M. R., et al. 2003a, ApJ, 592, 819
———. 2003b, AJ, 125, 2348
———. 2005, AJ, 129, 2562
Borch, A., et al. 2006, A&A, 453, 869
Carlberg, R. G., Yee, H. K. C., Ellingson, E., Abraham, R., Gravel, P., Morris, S., & Pritchet, C. J. 1996, ApJ, 462, 32
Cohn, J. D., Evrard, A. E., White, M., Croton, D., & Ellingson, E. 2007, MNRAS, in press (arXiv: 0706.0211)
Coil, A. L., et al. 2006, ApJ, 638, 668
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Colless, M., et al. 2001, MNRAS, 328, 1039
Collister, A. A., & Lahav, O. 2005, MNRAS, 361, 415
Cooper, M. C., Newman, J. A., Madgwick, D. S., Gerke, B. F., Yan, R., & Davis, M. 2005, ApJ, 634, 833
Cooray, A. 2006, MNRAS, 365, 842
Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1
Crook, A. C., Huchra, J. P., Martimbeau, N., Masters, K. L., Jarrett, T., & Macri, L. M. 2007, ApJ, 655, 790
Croton, D. J., et al. 2006, MNRAS, 365, 11
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H. G., Buta, R. J., Paturel, G., & Fouque, P. 1991, Third Reference Catalogue of Bright Galaxies (Heidelberg: Springer)

Einasto, J., et al. 2007, A&A, 462, 811
Eisenstein, D. J., & Hu W. 1998, ApJ, 496, 605
Eke, V. R., et al. 2004, MNRAS, 348, 866
Geller, M. J., & Huchra, J. P. 1983, ApJS, 52, 61
Gerke, B. F., et al. 2005, ApJ, 625, 6
Gladders, M. D., & Yee, H. K. C. 2005, ApJS, 157, 1
Gonzalez, A. H., Zaritsky, D., Dalcanton, J. J., & Nelson, A. 2001, ApJS, 137, 117
Goto, T. 2005, MNRAS, 359, 1415
Goto, T., et al. 2002, AJ, 123, 1807
Jing, Y. P., Mo, H. J., & Börner, G. 1998, ApJ, 494, 1
Kang, X., Jing, Y. P., Mo, H. J., & Börner, G. 2005, ApJ, 631, 21
Katz, N., Weinberg, D. H., & Hernquist, L. 1996, ApJS, 105, 19
Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201
Kauffmann, G., White, S. D. M., Heckman, T. M., Menard, B., Brinchmann, J., Charlot, S., Tremonti, C., & Brinkmann, J. 2004, MNRAS, 353, 713
Kim, R. J. S., et al. 2002, AJ, 123, 20
Koester, B. P., et al. 2007, ApJ, 660, 239
Kroupa, P. 2001, MNRAS, 322, 231
Lee, B. C., et al. 2004, AJ, 127, 1811
Li, C., Jing, Y. P., Kauffmann, G., Börner G, Kang, X., & Wang, L. 2007, MNRAS, 376, 984
Li, C., Kauffmann, G., Jing, Y. P., White, S. D. M., Börner, G., & Cheng, F. Z. 2006, MNRAS, 368, 21
Maccio, A. V., Dutton, A. A., van den Bosch, F. C., Moore, B., Potter, D., & Stadel, J. 2007, MNRAS, 378, 55
Merchán M., & Zandivarez, A. 2002, MNRAS, 335, 216
———. 2005, ApJ, 630, 759
Miller, C. J., et al. 2005, AJ, 130, 968
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Peacock, J. A., & Smith, R. E. 2000, MNRAS, 318, 1144
Pearce, F. R., Thomas, P. A., Couchman, H. M. P., & Edge, A. C. 2000, MNRAS, 317, 1029
Popesso, P., Böhringer, H., Brinkmann, J., Voges, W., & York, D. G. 2004, A&A, 423, 449
Robotham, A., Wallace, C., Phillipps, S., De Propris, R. 2006, ApJ, 652, 1077
Rozo, E., Wechsler, R. H., Koester, B. P., Evrard, A. E., McKay, T. A. 2007, ApJ, submitted (astro-ph/0703574)

Saunders, W., et al. 2000, MNRAS, 317, 55

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Scranton, R. 2003, MNRAS, 339, 410

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

Somerville, R. S., & Primack, J. R. 1999, MNRAS, 310, 1087

Spergel, D. N., et al. 2007, ApJS, 170, 377

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., et al. 2005, Nature, 435, 629

Stoughton, C., et al. 2002, AJ, 123, 485

Tago, E., et al. 2006, Astron. Nachr., 327, 365

Tinker, J. L., Weinberg, D. H., Zheng, Z., & Zehavi, I. 2005, ApJ, 631, 41

Tucker, D. L., et al. 2000, ApJS, 130, 237

Vale, A., & Ostriker, J. P. 2006, MNRAS, 371, 1173

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

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

van den Bosch, F. C., Weinmann, S. M., Yang, X., Mo, H. J., Li, C., & Jing, Y. P. 2005, MNRAS, 361, 1203

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

van den Bosch, F. C., Yang, X., Mo, H. J., Weinmann, S. M., Maccio, A., More, S., Cacciato, M., & Skibba, R. 2007, MNRAS, 376, 841

Warren, M. S., Abazajian, K., Holz, D. E., & Teodoro, L. 2006, ApJ, 646, 881

Weinmann, S. M., van den Bosch, F. C., Yang, X., & Mo, H. J. 2006a, MNRAS, 366, 2

Weinmann, S. M., van den Bosch, F. C., Yang, X., Mo, H. J., Croton, D. J., & Moore, B. 2006b, MNRAS, 372, 1161

White, S. D. M., & Frenk, C. 1991, ApJ, 379, 52

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

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

Yang, X., Mo, H. J., Jing, Y. P., & van den Bosch, F. C. 2005a, MNRAS, 358, 217

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

Yang, X., Mo, H. J., & van den Bosch, F. C. 2006, ApJ, 638, L55

Yang, X., Mo, H. J., van den Bosch, F. C., & Jing, Y. P. 2005b, MNRAS, 356, 1293

———. 2005c, MNRAS, 357, 608

Yang, X., Mo, H. J., van den Bosch, F. C., Weinmann, S. M., Li, C., & Jing, Y. P. 2005d, MNRAS, 362, 711

York, D. G., et al. 2000, AJ, 120, 1579

Zandivarez, A., Martinez, H. J., & Merchan, M. E. 2006, ApJ, 650, 137

Zehavi, I., et al. 2002, ApJ, 571, 172

Zheng, Z., et al. 2005, ApJ, 633, 791