Clustering of K-band selected local galaxies [*]

Bin Ma[1], Ke-Lai Meng[1], Jun Pan[1], Jia-Sheng Huang[2] and Long-Long Feng[1,3]

[1] Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, China; *mabin@pmo.ac.cn*
[2] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[3] Department of Physics, Texas A&M University, College Station, TX 77843, USA

Abstract We present a detailed clustering analysis of a large K-band selected local galaxy sample, which is constructed from the 2MASS and the SDSS and consists of 82 486 galaxies with $10 < K < 13.5$ and $0.01 < z < 0.1$. The two-point correlation function of the magnitude-limited sample in real space at small scales is well described by the power law $\xi(r) = (r/6.44 \pm 0.23)^{-1.81\pm0.02}$. The pairwise velocity dispersion is derived from the anisotropic two-point correlation function and we find the dispersion $\sigma_{12} = 685 \pm 17 \,\mathrm{km\,s^{-1}}$ if scale invariance is assumed, which is larger than values measured in the optical bands of the selected galaxy samples. We further investigate the dependence of the two-point correlation function and the σ_{12} on the $g-r$ color and the K-band luminosity, and obtain similar results to previous works in optical bands. Comparing a mock galaxy sample with our real data indicates that the semi-analytical model cannot mimic the σ_{12} in observation, although it can approximate the two-point correlation function within measurement uncertainties.

Key words: galaxies: statistics — infrared: galaxies — cosmology: large-scale structure of universe

1 INTRODUCTION

The track of galaxy formation and evolution is still one of the pivotal but intricate problems in modern astrophysical research. While there are large portions of prominent results mainly derived in optical bands, it is worth of addressing that observations of galaxies at distinct wavebands actually depict galaxies' intrinsic properties in different aspects, which is essential to our attempt at an unbiased understanding of the formation and evolution of galaxies. The Near-Infrared (NIR) observations are of particular interest. A K-band ($2.15\,\mu$m) selected galaxy sample has several advantages in galactic research. The K-correction in the K-band is a relatively small and better understood quantity compared with the optical K-correction. A galaxy's K-band light is $5 - 10$ times less sensitive to dust and stellar populations, and it is independent of its galaxy spectral type at $z < 1$ (e.g. Mobasher et al. 1986; Cowie et al. 1994). Therefore, K-band observations enable a much more homogeneous sampling of the galaxy types. It is even advantageous that the K-band luminosity is tightly correlated with the galaxy's stellar mass, so that the analysis at the K-band is a better probe of galaxy properties and relevant evolutionary history related to the stellar component than those measurements in the B- and r-bands.

Studying clustering strength using the two-point correlation function (2PCF) of galaxies in different redshift bins is one of the most straightforward and effective methods of exploring how galaxy properties are related to the underlying dark matter distribution and henceforth to the macro history of galactic formation and evolution. For example, the clustering of galaxies selected at $3.6\,\mu$m with median redshift $z \sim 0.5$ offered by the *Spitzer* Wide-area Infrared Extragalactic (SWIRE) survey is directly compared to the results of local galaxy samples selected at the K-band centered at $2.15\,\mu$m (Waddington et al. 2007). Unfortunately, in contrast with prolific photometric galaxy surveys available in NIR bands, spectroscopic information about these galaxies lags behind. It is not strange to see that the majority of clustering analysis of NIR galaxies resorts to angular correlation functions $\omega(\theta)$ (e.g. Baugh et al. 1996; McCracken et al. 2000; Roche et al. 2003; Waddington et al. 2007). The spatial correlation function $\xi(r)$ is only occasionally applicable (Carlberg et al. 1997).

In principle, one can invert the angular 2PCF to the real space spatial correlation function through Limber's equation (see p. 194 of Peebles 1980). The problem is that the precision of the inversion relies significantly on the radial selection function of galaxies and the assumptions of the small angle approximation and a power law $\xi(r)$, which introduces apparent uncertainties to its interpretation (Bernardeau et al. 2002), e.g. setting up the empirical evolution model of the real space two-point correlation function by $\xi(r, z) = (r/r_0)^{-\gamma} \times (1 + z)^{\gamma - 3 - \epsilon}$ (Phillipps et al. 1978).

Thanks to the Two-Micron All Sky Survey Extended Source Catalog (2MASS XSC Jarrett et al. 2000) and the Sloan Digital Sky Survey (SDSS York et al. 2000), we are able to construct a large K-band selected local galaxy sample to measure spatial two point correlation functions, and thus to build a clustering analysis of local NIR galaxies. By combining two-point correlation functions in both the redshift and real spaces, we are also able to derive quantities that describe the galaxies' peculiar velocities which are largely ignored in characterizing clustering of galaxies. Actually, to appropriately describe galaxy distributions, we should work in the phase space supported by both of the position and the peculiar velocity, especially in the nonlinear regime. Thus, measurement of the pairwise peculiar velocity dispersion is robust, providing important statistics to identify galaxy populations and test galaxy formation models (e.g. Zhao et al. 2002; Yoshikawa et al. 2003; Jing & Börner 2004; Slosar et al. 2006; Li et al. 2007).

A brief summary of the sample used is presented in Section 2, as well as the estimation method of 2PCF, including technical issues such as random sample construction and correction to the fiber collisions. In Section 3, we present measurements of the full, flux-limited sample in both redshift- and real-spaces, together with the redshift-space distortion parameter and the pairwise velocity dispersion. In Section 4, the full sample is divided into different sub-samples according to their color and luminosity to probe the clustering results as functions of these properties.

Throughout this paper, galaxy distances are obtained from redshifts assuming a cosmology with $\Omega_m = 0.3, \Omega_\Lambda = 0.7$, quoted in units of h^{-1} Mpc. Absolute magnitudes quoted for galaxies assume $h = 1$ to avoid the $-5 \log_{10} h$ factor.

2 SAMPLES AND ESTIMATION PROCEDURE

2.1 The K-band Sample of Local Galaxies

Our K-band galaxy sample is selected from the 2MASS XSC. Redshifts for this sample are obtained from the SDSS. The two catalogs actually have been combined by Blanton et al. (2005) into the New York University Value-Added Galaxy Catalog (NYU-VAGC) [1] to form a local redshift sample (mostly below $z \sim 0.3$) [2], with a coverage of $9938\,\mathrm{deg}^2$ for photometric imaging and $6750\,\mathrm{deg}^2$ for spectroscopic observation.

The completeness of the cross-matched 2MASS+SDSS Catalog has been extensively discussed (e.g. Bell et al. 2003). Following these practices, we select the galaxies by the extinction-corrected Kron magnitudes in the range $10 \leq K \leq 13.5$, and redshifts in the range $0.01 \leq z \leq 0.1$. In the apparent

---

[1] *http://sdss.physics.nyu.edu/vagc*

[2] The version of the NYU-VAGC used in this paper is the SDSS DR6 (Adelman-McCarthy et al. 2008).

Table 1 K-selected Samples including both Red and Blue Populations

Label	N_{gal}	M^*	α	$r_0(\,h^{-1}\,\mathrm{Mpc})$	γ	β
Main	82 486	−23.38	−0.94	6.44 ± 0.23	1.81 ± 0.02	0.39 ± 0.17
Red	46 689	−23.31	−0.48	7.62 ± 0.32	1.87 ± 0.03	0.35 ± 0.20
Blue	35 797	−23.08	−1.11	4.92 ± 0.14	1.67 ± 0.02	0.50 ± 0.15

Notes: All samples are in the range of $10 < K < 13.5$ and $0.01 < z < 0.1$. M^* and α are fitted parameters of the Schechter luminosity function (Schechter 1976). r_0 and γ are obtained from the fit to the projected 2PCF $w_{\mathrm{p}}(\sigma)$. The numbers of galaxies quoted here include those that lack redshifts due to fiber collisions.
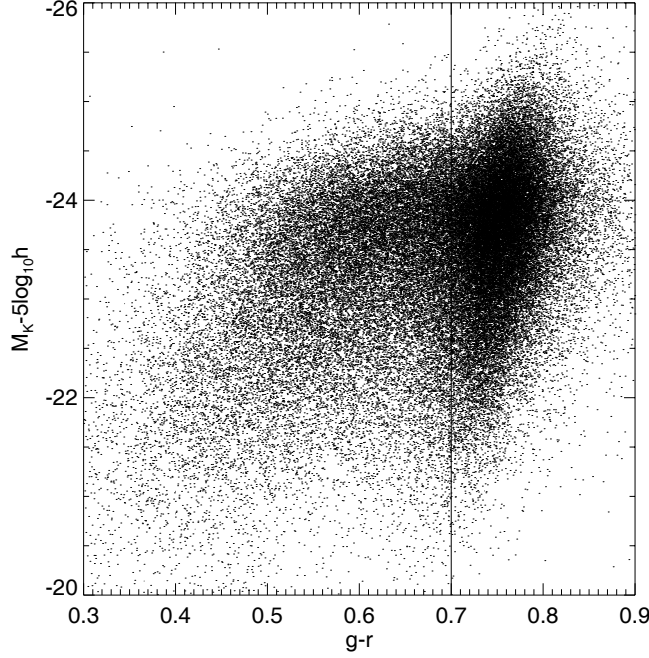
Fig. 1 K-corrected $g - r$ color vs. K-band absolute magnitude for all galaxies. The vertical line demarcates a simple cut at $g - r = 0.7$.

magnitude limits adopted here, the bright end is chosen to avoid incompleteness due to the large angular sizes of galaxies, while the faint end is used to match the magnitude limit of 2MASS. We also use a low redshift cut which rejects galaxies with redshifts seriously affected by Hubble flow. The final flux-limited sample, our main sample, accumulates a total of 78 339 galaxies with redshifts, with a median redshift of ~ 0.06.

The K-selected sample is further divided into two sub-samples according to their SDSS color $(g - r)$: the blue sample with $(g - r) < 0.7$ and red sample with $(g - r) > 0.7$ (Fig. 1 and Table 1). Both red and blue samples have similar space densities, but galaxies in the red sample are systematically more luminous and therefore mainly show higher redshifts: the median redshift of red galaxies is $z_{\mathrm{med}} \sim 0.065$ whereas for the blue ones $z_{\mathrm{med}} \sim 0.054$.

A set of volume-limited sub-samples within different absolute magnitude bins is also constructed for analysis (Table 2).

Table 2 Volume-limited Sub-samples and Their 2PCF Parameters

Abs. Mag.	Redshift range	$N_{\rm gal}$	$\bar{n}$	r_0	γ
$-25 \sim -24$	$0.033 - 0.100$	14 152	0.87	$7.17 \pm 0.30\ (6.93 \pm 1.10)$	$1.87 \pm 0.03\ (1.82 \pm 0.07)$
$-24 \sim -23$	$0.021 - 0.067$	19 714	3.83	$5.52 \pm 0.28\ (4.97 \pm 0.23)$	$1.83 \pm 0.03\ (1.87 \pm 0.03)$
$-23 \sim -22$	$0.013 - 0.042$	8969	7.08	$5.23 \pm 0.65\ (4.50 \pm 0.24)$	$1.83 \pm 0.04\ (1.86 \pm 0.04)$
$-22 \sim -21$	$0.010 - 0.026$	2232	7.19	$4.46 \pm 0.57\ (4.79 \pm 0.29)$	$1.86 \pm 0.10\ (1.92 \pm 0.04)$

Notes: $\bar{n}$ is the mean number density in units of $10^{-3}\ h^3\ {\rm Mpc}^{-3}$, and r_0 is in units of h^{-1} Mpc. All galaxy samples are in the range $10 < K < 13.5$, $0.01 < z < 0.1$. r_0 and γ are obtained from $w_{\rm p}(\sigma)$, the numbers in parentheses are values of corresponding luminosity sub-samples of the mock galaxy catalog.

2.2 Mock Galaxy Catalog

We also compare our sample with a mock galaxy catalog to scout for the performance of semi-analytical models for hierarchical galaxy formation (see Baugh 2006, and references there in). The mock catalog is derived from a high-resolution pure dark matter cosmological simulation with 512^3 particles in a box of size $100\,h^{-1}$ Mpc run by the GADGET2 code (Springel 2005). The fundamental cosmological parameters for the simulation are set as $\Omega_\Lambda = 0.762$, $\Omega_m = 0.238$, $h = 0.73$, $\sigma_8 = 0.74$, $n = 1$, and the rest of the simulation parameters are the same as those in the pure dark matter run by Lin et al. (2006), e.g. the particle mass is $5.5 \times 10^8 h^{-1} M_\odot$ and the soft length is $4.5\,h^{-1}$ kpc. We generate the mock galaxies at $z = 0.05$ from the output of the simulation by the semi-analytical model of Kang et al. (2005). Luminosities are the only derived property for the mock galaxies, thus we divide the mock galaxies into a set of luminosity bins and study galaxy and dark halo properties in each bin.

2.3 Estimation of the Two-point Correlation Functions

Two types of 2PCFs for this sample are measured to capture the clustering patterns: one is the isotropic $\xi(s)$ with s denoting the separation of a pair of galaxies in redshift space, and the other is the two-dimensional function $\xi(\sigma, \pi = \sqrt{s^2 - \sigma^2})$ with π indicating the separation of a pair of galaxies parallel to the line-of-sight and σ being the separation perpendicular to the line-of-sight. The latter is mainly utilized to obtain the redshift distortion free function, the projected 2PCF $w_{\rm p}(\sigma) = \int \xi(\sigma, \pi) {\rm d}\pi$, and consequently, the real space function $\xi(r)$ after proper inversion (e.g. Davis & Peebles 1983; Hawkins et al. 2003).

The 2PCFs are measured using the estimator of Landy & Szalay (1993),

$$\xi = \frac{DD - 2DR + RR}{RR} , \qquad (1)$$

in which DD is the normalized number of weighted galaxy-galaxy pairs at given separations, RR is the normalized number of random-random pairs within the same separation in the random catalog and DR is the normalized number of weighted galaxy-random pairs. In general, the scale s of $\xi(s)$ is binned uniformly in a logarithmic scale, and for $\xi(\sigma, \pi)$, σ and π are binned in a linear scale.

To proceed with the estimation of Equation (1), an auxiliary sample of completely random points in the exact same geometric window as the galaxy sample is a prerequisite. The random samples should have the same redshift, magnitude and mask constraints as the real data, with a smooth selection function matching the $N(z)$ of the real data. The luminosity functions of the flux-limited samples, used to generate the selection function, are computed with the STY method (Sandage et al. 1979) in the form of the Schechter function (Table 1). We generate a random sample ten times larger than the observed K-selected sample, and the random samples for the red and blue samples are 15 times larger than the observed ones.

A weight is assigned to each galaxy and a random point according to their redshift and angular position to minimize the variance in estimating ξ (Efstathiou 1988; Hamilton 1993),

$$\omega_i = \frac{1}{1 + 4\pi\, n(z)\Phi_i\, J_3(s)} \,, \tag{2}$$

where Φ_i is the selection function at the location of the ith galaxy, $n(z)$ is the mean number density, and $J_3(s) = \int_0^s \xi(s)s^2 ds$. The $J_3(s)$ is computed using a power-law $\xi(s)$ with correlation length $s_0 = 12\,h^{-1}\,\mathrm{Mpc}$ and $\gamma_0 = 1$. We find that the difference between our estimate and that using a raw measured $\xi(s)$ is negligible, in agreement with the conclusion of Hawkins et al. (2003). To normalize the pair counts properly, we assure that the sum of weights of the random catalog equals the sum of weights of the real galaxy catalog, both of which are functions of scale.

We need to correct the incompleteness in the spectroscopic sample due to the effect of colliding fiber constraints. The design of the SDSS instrument means that fibers cannot be placed closer than 55 arcsec on the same tile. Members of a close pair of galaxies cannot be targeted in a single fiber configuration so that the survey misses a large fraction of close pairs. Because 2PCF will be systematically underestimated without taking account of the fiber collision effect, several methods are developed to correct this bias (Zehavi et al. 2002; Hawkins et al. 2003). We adopt the method of Zehavi et al. (2002) by assigning the redshift of the observed galaxy in a pair to the pair member whose redshift was absent. Then, we obtain the "collision corrected" redshift. At large scales, where both members of the pair contribute to the same separation bin, this method is equivalent to double weighting. We argue that this method should perform better on small scales because it retains information about the known angular positions. Zehavi et al. (2005) showed that this method is an adequate treatment: residual systematics for the redshift space correlation function were considerably smaller than the statistical errors, and this was even more true for the projected 2PCF $w_\mathrm{p}(\sigma)$.

The covariance matrix is computed with the jack-knife technique (Lupton 1993; Zehavi et al. 2002), where the galaxy sample is split into thirty separate regions of approximately equal sky area, and then we perform the analysis thirty times, with each time leaving a different region out of the calculation. The estimated statistical covariance of 2PCF measured in two bins of i and j is then

$$\mathrm{Cov}(\xi_i, \xi_j) = \frac{N-1}{N} \sum_{\ell=1}^{N} (\xi_{i,\ell} - \overline{\xi_i})(\xi_{j,\ell} - \overline{\xi_j}) \,, \tag{3}$$

in which N is the number of jack-knife sub-samples.

3 CLUSTERING OF THE MAIN SAMPLE

3.1 2PCF in Redshift Space

The first one we calculated is the 2PCF $\xi(s)$ for the main sample from $s = 0.01\,h^{-1}\,\mathrm{Mpc}$ to $s = 100\,h^{-1}\,\mathrm{Mpc}$ in redshift space. s is divided equally into logarithmic bins of width $\Delta \log_{10}(s/\,h^{-1}\,\mathrm{Mpc}) = 0.125$. $\xi(s)$ is not a single power law at all scales (Fig. 2), but rather

$$\xi(s) = \left(\frac{s}{s_0}\right)^{-\gamma_s}, \quad (s_0, \gamma_s) = \begin{array}{ll} (11.69 \pm 0.89, 0.98 \pm 0.02), & \text{if } s \in (0.1, 3) \,; \\ (7.64 \pm 0.41, 1.60 \pm 0.05), & \text{if } s \in (3, 10) \,. \end{array} \tag{4}$$

It is known that $\xi(s)$ also consists of the contribution from the galaxies' peculiar velocities causing the redshift distortion, in addition to the true spatial correlation of galaxies (Kaiser 1987; Hamilton 1998). Thus, we need to break the degeneracy between the spatial clustering and the velocity correlation before making direct comparisons of $\xi(s)$ in different redshift bins.
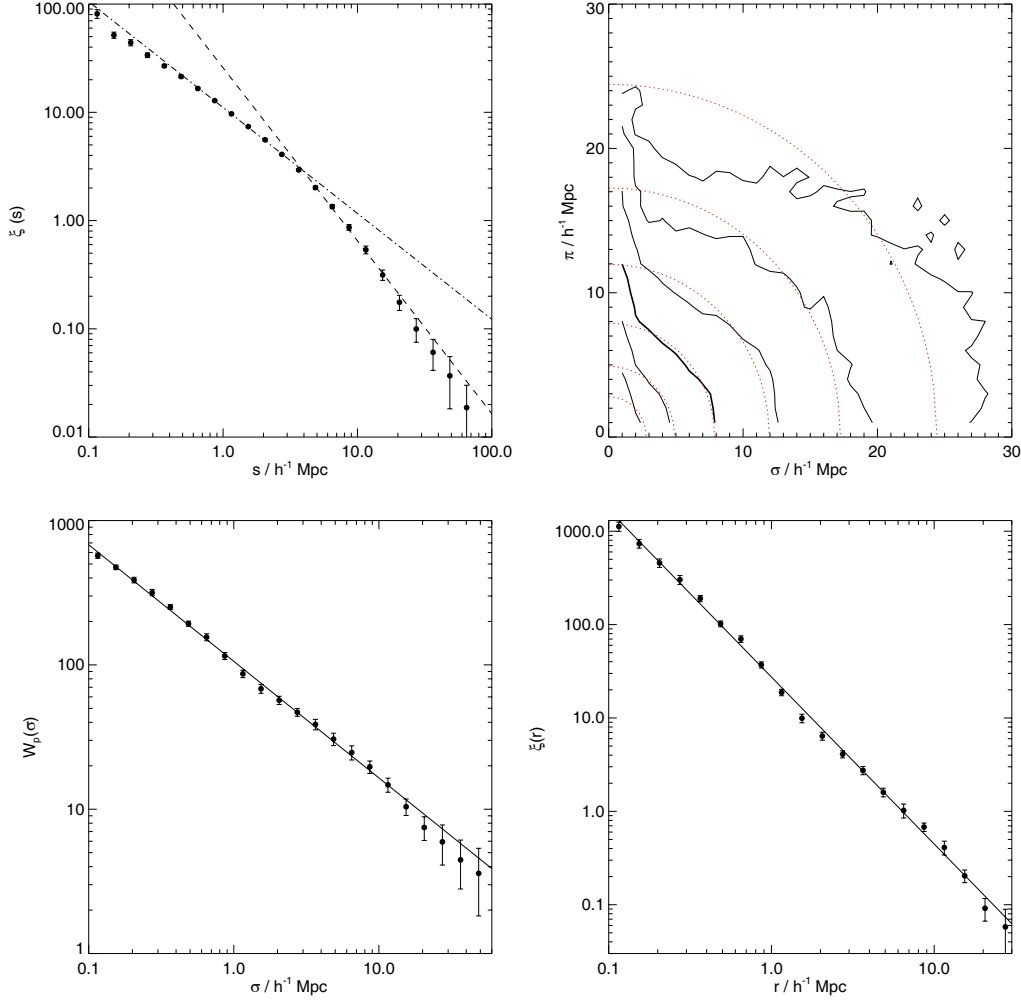
Fig. 2 2PCFs of the main sample. *Top-left*: ξ in redshift space, the dot-dashed line is the power law at scales $0.1 < s < 3\,h^{-1}\,\mathrm{Mpc}$ and the dashed line is the fit at scales $3 < s < 10\,h^{-1}\,\mathrm{Mpc}$ (Eq. (4)). *Top-right*: $\xi(\sigma, \pi)$ contours, σ and π are separations perpendicular and parallel to the line of sight respectively, solid contours are measurements at levels of $4, 2, 1, 0.5, 0.25, 0.125$ outwards and dotted lines are the isotropic 2PCF expected in the absence of redshift-space distortion. *Bottom-left*: projected 2PCF $w_{\mathrm{p}}(\sigma)$ (solid points), the solid line is the power law fit. *Bottom-right*: 2PCF $\xi(r)$ in real space inverted from w_{p} through Eq. (7), the solid line is the power law transformed from the fit to the w_{p} on the left. Error bars in these plots are the 1σ scatters of jackknife sub-samples.

3.2 The Projected 2PCF

The galaxy's peculiar motion only causes drifting of the radial position. To minimize this effect, we can calculate the correlation function as a function of σ and π, where σ is perpendicular to the line-of-sight and π is parallel to the line-of-sight. Then, the projection of $\xi(\sigma, \pi)$ onto the σ plane is independent of redshift space distortion and gives information about the real space correlation function.

The effect of redshift distortion is clearly seen in the $\xi(\sigma, \pi)$ contour of the main sample in Figure 2: the contours are elongated along the line-of-sight direction at small σ separation, exhibiting the phe-

Table 3 Clustering of galaxies selected at different bands. The r-band (SDSS) result comes from Zehavi et al. (2002, 2005) and the b_J-band (2dFGRS) is given by Hawkins et al. (2003).

Band	N_{gal}	r_0 (h^{-1} Mpc)	γ	σ_{12} (km s^{-1})
Ks	82 486	6.44 ± 0.23	1.81 ± 0.02	685 ± 17
r	154 014	5.59 ± 0.11	1.84 ± 0.01	640 ± 60
b_J	165 659	4.95 ± 0.27	1.72 ± 0.04	506 ± 52

nomenon of fingers-of-God by the random pairwise velocity; at large σ scales, the contours are squashed in the π direction due to gravitational coherent inflow (Hawkins et al. 2003).

Integrating the anisotropic $\xi(\sigma, \pi)$ over π gives the projected 2PCF,

$$w_{\text{p}}(\sigma) = \int_{-\pi_{\max}}^{+\pi_{\max}} \xi(\sigma, \pi) d\pi = \sum_i \xi(\sigma, \pi_i) \Delta \pi_i , \tag{5}$$

which has the practical upper limit $\pi_{\max} = 50 \, h^{-1}$ Mpc. We found that there is little difference if a larger cutoff is used.

w_{p} is related to the 2PCF in real space through the Abel transform (Davis & Peebles 1983)

$$w_{\text{p}}(\sigma) = 2 \int_{\sigma}^{\infty} \frac{r \xi(r) \, dr}{\sqrt{r^2 - \sigma^2}} , \tag{6}$$

where the inversion renders $\xi(r)$ assuming a step function-like $w_{\text{p}}(\sigma) = w_i$ at each bin centered at σ_i

$$\xi(\sigma_i) = \frac{1}{\pi} \sum_{j \geq i} \frac{w_{j+1} - w_j}{\sigma_{j+1} - \sigma_j} \ln \left(\frac{\sigma_{j+1} + \sqrt{\sigma_{j+1}^2 - \sigma_i^2}}{\sigma_j + \sqrt{\sigma_j^2 - \sigma_i^2}} \right) \tag{7}$$

for $r = \sigma_i$ (Saunders et al. 1992). Mathematically, the inversion is not stable but works well in practice in $r < 30 \, h^{-1}$ Mpc. If we simply assume $\xi(r) = (r/r_0)^{-\gamma}$, the integral in Equation (6) can be done analytically, yielding

$$w_{\text{p}} = A\sigma^{1-\gamma}, \text{ with } A = r_0^{\gamma} \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{\gamma-1}{2}\right) / \Gamma\left(\frac{\gamma}{2}\right), \tag{8}$$

where Γ is the Gamma function. It is true that w_{p} for our main sample is a power law function at small scales (Fig. 2); the best-fit parameters are $\gamma = 1.81 \pm 0.02$ and $r_0 = 6.44 \pm 0.23$ in the regime of $0.1 \, h^{-1}$ Mpc $< \sigma < 18 \, h^{-1}$ Mpc.

Table 3 lists the results of our K-selected sample and those selected in the other bands for comparison: b_J of 2dFGRS and r of SDSS. We conclude that the correlation function will have a larger amplitude and steeper slope if the galaxies are selected at a longer wavelength band.

3.3 The Pairwise Velocity Dispersion

Currently, there is no precise theory about what causes the full scale redshift distortion (Scoccimarro 2004). However, we can still approach this topic with reasonable assumptions. Intuitively, redshift distortion can be approximated by a certain convolution of two components dominated in different regimes. Coherent infall is responsible for the clustering enhancement at large scales while the smearing of correlation strength at small scales is attributed to random motions.

Kaiser (1987) found that at large scales, the boost to the power spectrum by the peculiar velocities takes a particularly simple form, which was later completed and translated to real space by Hamilton (1992),

$$\xi'(\sigma, \pi) = \xi_0(s) P_0(\mu) + \xi_2(s) P_2(\mu) + \xi_4(s) P_4(\mu) , \tag{9}$$

where $P_\ell(\mu)$ are Legendre polynomials, $\mu = \cos(\theta)$ with θ being the angle between r and π. Assuming $\xi = (r/r_0)^{-\gamma}$ then renders

$$\xi_0(s) = \xi(s) = \left(1 + \frac{2\beta}{3} + \frac{\beta^2}{5}\right)\xi(r),$$

$$\xi_2(s) = \left(\frac{4\beta}{3} + \frac{4\beta^2}{7}\right)\left(\frac{\gamma}{\gamma - 3}\right)\xi(r), \tag{10}$$

$$\xi_4(s) = \frac{8\beta^2}{35}\left(\frac{\gamma(2 + \gamma)}{(3 - \gamma)(5 - \gamma)}\right)\xi(r),$$

where s and r have the corresponding values, β is the linear redshift distortion parameter, $\beta \approx \Omega_0^{0.6}/b$, and b is the linear bias parameter. The first equation is independent of the form of $\xi(r)$.

To incorporate effects of random motion, the anisotropic 2PCF in redshift space is then approximated by the convolution of $\xi'(\sigma, \pi)$ by Equation (9) with the distribution function of the pairwise velocity $f(v_{12})$ (c.f. Peebles 1993),

$$\xi(\sigma, \pi) = \int_{-\infty}^{+\infty} \xi'(\sigma, \pi - \frac{v_{12}}{H_0})f(v_{12})dv_{12}, \tag{11}$$

and, in general, $f(v_{12})$ is assumed to be an exponential distribution of dispersion σ_{12}

$$f(v_{12}) = \frac{1}{\sigma_{12}\sqrt{2}}\exp\left(-\frac{\sqrt{2}v_{12}}{\sigma_{12}}\right). \tag{12}$$

After measuring $\xi(s)$ and $\xi(r)$ at large scales, we can easily obtain the redshift distortion parameter β via the first equation in Equation (10). Then, we combine Equations (9) – (12) to fit the $\xi(\sigma, \pi)$ data grid to determine the other model parameters. However, there are implicit assumptions in the prescription: (1) the linear bias parameter b is forced to be scale-independent; (2) the pairwise velocity dispersion σ_{12} is presumed invariant with respect to the separation π along the line-of-sight but could be a function of the perpendicular separation σ. Therefore, the resulting σ_{12} in the model is not the actual true pairwise velocity dispersion. To have a meaningful comparison with simulations, we need to estimate the σ_{12} from the 2PCFs of the simulation data in the same way as the galaxy sample.

Figure 3 shows the ratio of the redshift-space 2PCF to the real-space 2PCF, and the derived σ_{12} as a function of σ. $\xi(s)/\xi(r)$ increases with increasing scales, and becomes roughly constant in the range $4\,h^{-1}\,\mathrm{Mpc} < r < 13\,h^{-1}\,\mathrm{Mpc}$ as expected. However, the ratio again increases at larger scales. If we exclude the galaxies in the Sloan Great Wall region ($11^{\mathrm{h}} < \mathrm{R.A.} < 14^{\mathrm{h}}$), the $\xi(s)/\xi(r)$ ratio indeed does not show the up-shooting anymore. Nevertheless, because there are no appealing arguments about whether chopping off galaxies in the Sloan Great Wall guarantees fairness, we just pick up points in the scale range $(4, 13\,h^{-1}\,\mathrm{Mpc})$ to fit to a constant and get $\xi(s)/\xi(r) = 1.29 \pm 0.14$, or $\beta = 0.39 \pm 0.17$, considering the fact that the final σ_{12} is not at all sensitive to the β (Li et al. 2006a).

The σ_{12} as a function of projected separation σ shown in Figure 3 has the same classical shape as many other measurements (e.g. Jing & Börner 2004; Li et al. 2007): σ_{12} rises from $650\ \mathrm{km\,s^{-1}}$ up to $750\ \mathrm{km\,s^{-1}}$ as σ increases from $0.1\,h^{-1}\,\mathrm{Mpc}$ to $0.3\,h^{-1}\,\mathrm{Mpc}$, forms a plateau to $\sigma \sim 1\,h^{-1}\,\mathrm{Mpc}$, then drops down again at larger scales. If a scale independent σ_{12} is assumed, the best fitted $\sigma_{12} = 685 \pm 17\ \mathrm{km\,s^{-1}}$ within $0.1\,h^{-1}\,\mathrm{Mpc} < r < 18\,h^{-1}\,\mathrm{Mpc}$. Our σ_{12} is significantly higher than that of the b_J-band sample (2dFGRS, Hawkins et al. 2003) and slightly larger than the r-band sample (SDSS, Zehavi et al. 2002).

4 COLOR SUB-SAMPLES

Figure 4 displays the clustering dependence on the color of galaxies, from $\xi(s)$ and w_{p} as numerous works have already revealed that spatial clustering for red galaxies is much stronger than for blue galaxies (see also Table 1). The slope of the 2PCF is a direct indicator of the strength of the galaxy's dynamic
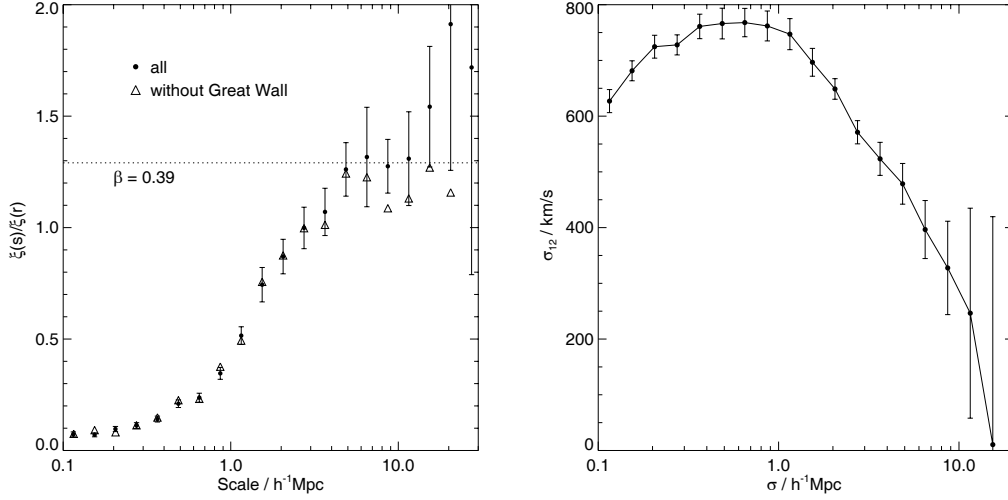
Fig. 3 *Left*: ratios of $\xi(s)/\xi(r)$ of the full sample (*solid circles*) in comparison with those of the sample without the Sloan Great Wall (*open triangles*); the dotted line corresponds to $\beta = 0.39$. *Right*: the derived pairwise velocity dispersion $\sigma_{12}(\sigma)$. Error bars are obtained by jackknife re-sampling.

non-linearity, so red galaxies with steeper 2PCFs are more harassed by gravitational turbulence at small scales or local structures than blue galaxies. In the scheme of halo models, the slope of 2PCF is determined by the percentage of contribution from the one-halo term and two-halo terms, rather than the mass of the halo in which the galaxy resides (Cooray & Sheth 2002). A steeper 2PCF at small scales contains more power from the one-halo term. If galaxies are simply central and satellites in dark halos, it leads to a conclusion that the satellite fraction in the red sub-sample is higher than in the blue sub-sample, although red galaxies are inclined to occupy more massive halos than blue galaxies. In fact, the interpretation is confirmed by the velocity information: $\sigma_{12}(\sigma)$ of blue galaxies is very flat and has a much lower amplitude than red galaxies over a wide range of projected separations, which is exactly what is observed in simulations when reducing the satellite fraction in mock galaxy samples (Slosar et al. 2006). We also note that the color dependence of σ_{12} is very similar to the simulation results of the old and young populations of galaxies demonstrated by Weinberg et al. (2004).

5 VOLUME-LIMITED SUB-SAMPLES AND COMPARISONS WITH OUR MOCK CATALOG

We construct four volume-limited sub-samples to investigate the luminosity dependence of clustering (Table 2). The main sample is divided to four absolute magnitude bins centered approximately from $M^* + 2$ to $M^* - 1$, where $M^* = -23.38$ is the characteristic luminosity of the Schechter function (Schechter 1976). The number density for galaxies in the lowest luminosity bin is eight times higher than that in the highest luminosity bin.

Figure 5 shows the projected correlation functions $w_{\rm p}(\sigma)$ of volume-limited samples, and Table 2 lists the parameters r_0 and γ for power law $\xi(r)$ models by fitting $w_{\rm p}(r_{\rm p})$ in the range $0.1\,h^{-1}\,{\rm Mpc} < \sigma < 13\,h^{-1}\,{\rm Mpc}$. The slopes of these sub-samples are roughly constant with error bars, indicating a weak dependence on luminosity; but the correlation length r_0 does increase with luminosity, thus being consistent with earlier studies (Norberg et al. 2001; Zehavi et al. 2005; Li et al. 2006b), and the $r_0 - L$ relation follows the same law as the relations in other bands and simulations (Fig. 5).

We also explore the relative bias factor b/b^* computed using the ratio of the $w_{\rm p}(2.7\,h^{-1}\,{\rm Mpc})$ of our four sub-samples to the $w_{\rm p}(2.7\,h^{-1}\,{\rm Mpc})$ of the $-24 < M_K < -23$ sub-sample ($L \approx L^*$). This
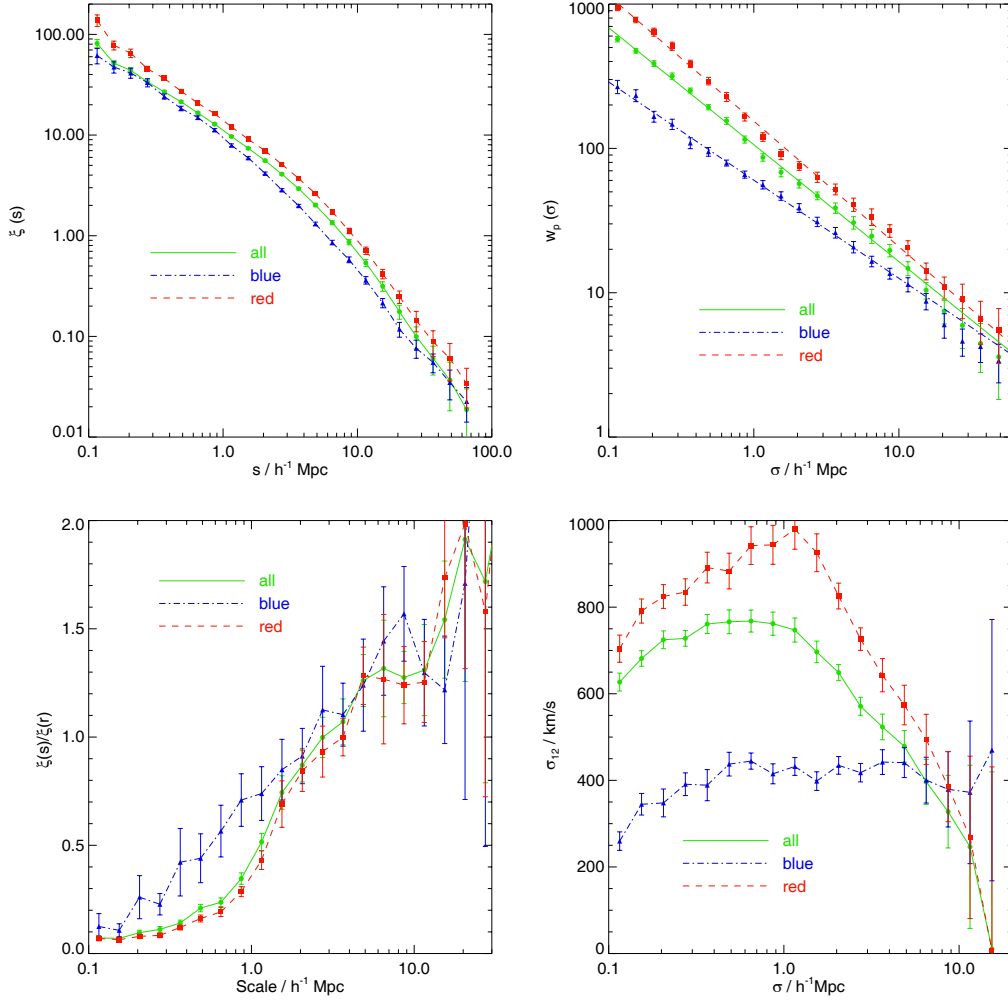
Fig. 4 $g - r$ color dependence of 2PCFs and σ_{12}. All error bars are derived from jackknife sub-samples.

fiducial separation of $2.7\,h^{-1}\,\mathrm{Mpc}$ is chosen because it is far from the extremely non-linear regime, but still small enough so that the correlation functions are very accurately measured for all sub-samples. In the bottom right panel of Figure 5, we compared our measured b/b^* to the models of Norberg et al. (2001), Tegmark et al. (2004) and Li et al. (2006b) together with data from simulations. Within the estimated uncertainties, the models of Norberg et al. (2001) and Tegmark et al. (2004) are basically consistent with our estimations, but the model of Li et al. (2006b) and the semi-analytical model fail to produce enough clustering power in the highest luminosity bin.

The luminosity dependence of σ_{12} is somewhat more complicated as already realized by Jing & Börner (2004) and Li et al. (2006a). It is pointed out that, at small separation, σ_{12} has a trough at L^*, becomes relatively flat at the lower luminosity end and increases rapidly at $L > L^*$. This has not been reproduced by any current halo models. Figure 6 shows that the σ_{12} curve of the L^* sub-sample at $\sigma < 1\,h^{-1}\,\mathrm{Mpc}$ is fairly flat and has a smaller amplitude than other luminosity sub-samples, implying a relatively smaller fraction of satellites are within the luminosity bin or the stellar mass range. Interestingly, the σ_{12} curves of the brightest sub-sample and the faintest sub-sample have very simi-
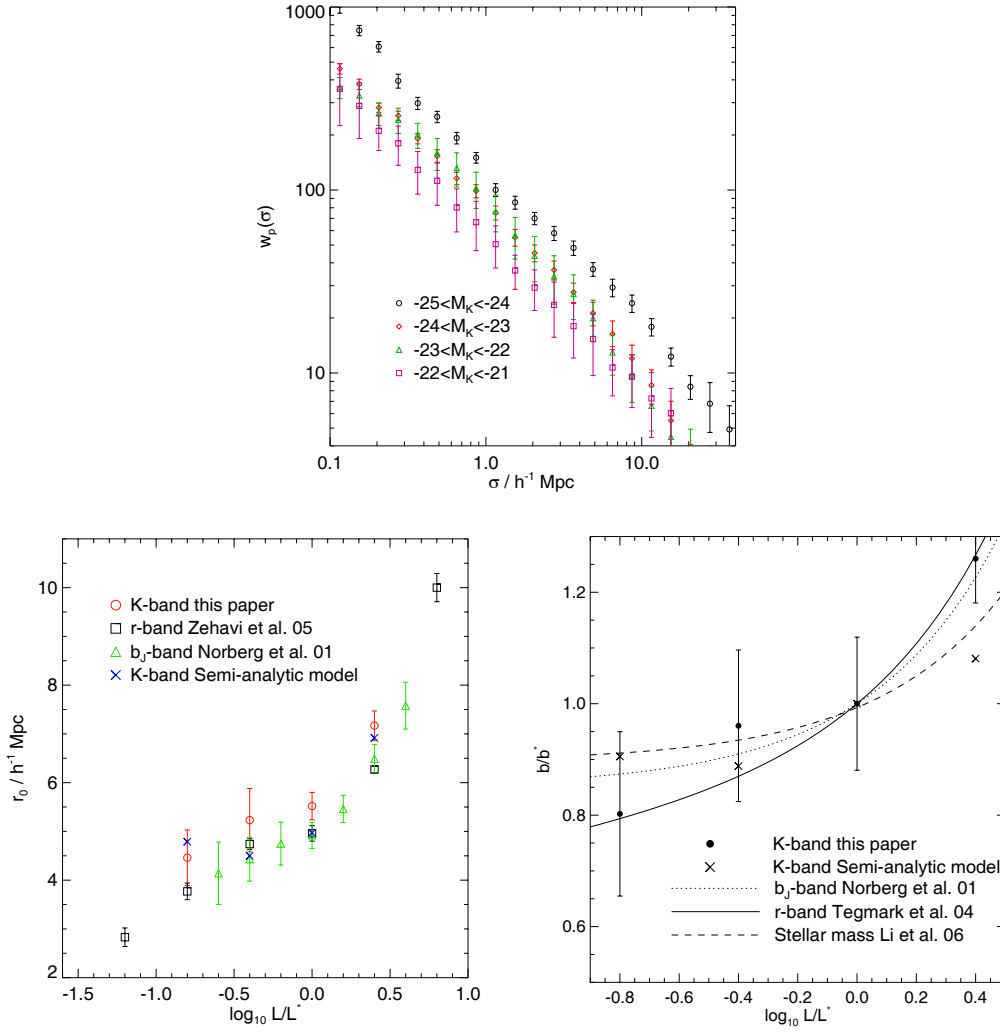
Fig. 5 *Top*: projected correlation functions $w_{\mathrm{p}}(\sigma)$ for the volume-limited sub-samples. *Bottom*: luminosity dependence of the comoving correlation lengths and the linear bias parameters of different galaxy samples and the mock galaxy sample. The relative bias factor b/b^* is defined by the amplitude of $w_{\mathrm{p}}(\sigma)$ measured at $\sigma = 2.7\,h^{-1}\,\mathrm{Mpc}$ divided by the same quantity of the $-24 < M_K < -23$ sub-sample ($L \approx L^*$); the dotted curve is at the b_J-band of the 2dFGRS survey $b/b^* = 0.85 + 0.15L/L^*$ (Norberg et al. 2001), the solid curve is from the SDSS r-band power spectrum $b/b^* = 0.85 + 0.15L/L^* - 0.04(M - M^*)$ (Tegmark et al. 2004); and the dashed curve is a fit to the $w_{\mathrm{p}}(\sigma)$ of stellar mass sub-samples $b/b^* = 0.896 + 0.097M/M^*$ (Li et al. 2006b) in which M and M^* refer to stellar mass.

lar shapes although their amplitudes differ slightly. It seems that the two sub-samples contain similar fractions of satellites although their occupied halos have very different masses, since the brightest sub-sample has a much bigger bias than the faintest sub-sample.

The semi-analytical model works well for galaxies with low K-band luminosity, but displays significant discrepancies at the high luminosity end despite the rough agreement in spatial clustering (Table 2). The case in the K-band is different from the discovery of Li et al. (2007) at the r-band where there are
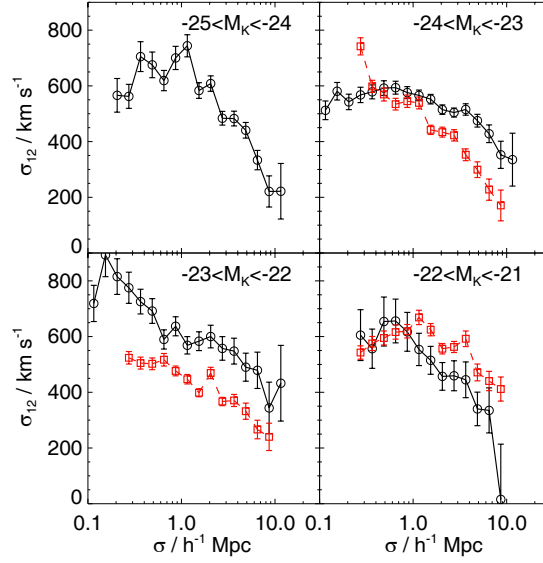
Fig. 6 σ_{12} of volume limited sub-samples (circles) against semi-analytical modeling (squares).

less differences between the semi-analytical model and the data at the lower luminosity end. However, the inconsistency of the semi-analytical model may be caused by the limited size of the simulation, leading to a deficiency in large mass halos that provide a sufficient number of galaxies of high stellar masses.

6 SUMMARY AND DISCUSSION

For furnishing the clustering evolution of NIR galaxies, a local galaxy catalog limited by K-band magnitudes of $10 < K < 13.5$ is generated by cross-matching the 2MASS data with the SDSS survey. Then, we carefully measured the 2PCFs and the pairwise velocity dispersions of the galaxy sample, as well as their color and luminosity dependence.

In the redshift space, the 2PCF $\xi(s)$ of the flux-limited sample complies better with the shallower power law at scales of $s <\sim 3\,h^{-1}$ Mpc than that at larger scales, but in real space the 2PCF $\xi(r)$ derived from the projected 2PCF. $w(\sigma)$ is perfectly approximated by a single power law with $r_0 = 6.44 \pm 0.23\,h^{-1}$ Mpc and $\gamma = 1.81 \pm 0.02$ for $r < 1 \sim 10\,h^{-1}$ Mpc (Fig. 2). The estimated correlation length supports the conclusion of Waddington et al. (2007) that the clustering of NIR galaxies evolves very slowly up to $z \sim 0.5$. The pairwise velocity dispersion at small scales is directly related to the spatial distribution in the halo (Slosar et al. 2006). Our σ_{12} of local NIR galaxies has a shape analogous to those found in simulations, showing a bump at scales around $1\,h^{-1}$ Mpc, but is larger than in optical bands, being 685 ± 17 km s^{-1} if assuming a scale invariant σ_{12} (Table 3). The phenomenon could be rooted in different luminosity functions at these bands, and it is likely that the distribution of K-band luminosities or stellar masses of satellite galaxies is more concentrated in the range defining our sample than that of central galaxies. Regrettably, we do not have a large NIR galaxy sample at high redshift, otherwise a comparison of σ_{12} would enable us to peer into the evolution of the positioning of NIR galaxies in their host halos, or the velocity biasing of galaxies relative to the dark matter.

The optical color dependence of the clustering of NIR galaxies is similar to the optical bands selected galaxies. Blue galaxies of $g - r < 0.7$ are much less clustered than red galaxies, and they display a very flat and low σ_{12} in the nonlinear regime as a consequences of having a smaller fraction of satellite

galaxies. In addition, the luminosity dependence of 2PCF and σ_{12} of our NIR galaxies is not different from that of optical galaxies in spite of the very different luminosity functions. As the K-band luminosity is tightly correlated with the stellar mass, it is somehow surprising to discover that the empirical formula of relative biasing by stellar mass in Li et al. (2006b) under-predicts the bias of the brightest volume-limited sample. Also, the transition of the shape and amplitude of σ_{12} at L^* might infer that the luminosity distribution of satellite galaxies in low mass halos is very different from that in high mass halos.

Examination of our galaxy sample against the mock galaxy sample reveals that the K-band luminosity dependence of the 2PCF can be approximately reproduced by semi-analytical modeling within measurement errors, but the σ_{12} of the mock one deviates from observation significantly in aspects of amplitude, shape and luminosity dependence, especially for bright galaxies. The peculiar velocity of a galaxy is a conundrum in galaxy formation models both in theories and simulations, as accurate modeling requires exquisite fabrication of galaxies in halos and corresponding evolution paths. It seems that, so far, there is still a long way to go.

References

Adelman-McCarthy, J. K., et al. 2008, ApJS, 175, 297

Baugh, C. M. 2006, Reports on Progress in Physics, 69, 3101

Baugh, C. M., Gardner, J. P., Frenk, C. S., & Sharples, R. M. 1996, MNRAS, 283, 15

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Phys. Rep., 367, 1

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

Carlberg, R. G., Cowie, L. L., Songaila, A., & Hu, E. M. 1997, ApJ, 484, 538

Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1

Cowie, L. L., Gardner, J. P., Hu, E. M., Songaila, A., Hodapp, K.-W., & Wainscoat, R. J. 1994, ApJ, 434, 114

Davis, M., & Peebles, P. J. E. 1983, ApJ, 267, 465

Efstathiou, G. 1988, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 297, Comets to Cosmology, ed. A. Lawrence, 312

Hamilton, A. J. S. 1992, ApJ, 385, 5

Hamilton, A. J. S. 1993, ApJ, 417, 19

Hamilton, A. J. S. 1998, in ASSL Vol. 231: The Evolving Universe, 185

Hawkins, E., et al. 2003, MNRAS, 346, 78

Jarrett, T. H., Chester, T., Cutri, R., Schneider, S., Skrutskie, M., & Huchra, J. P. 2000, AJ, 119, 2498

Jing, Y. P., & Börner, G. 2004, ApJ, 617, 782

Kaiser, N. 1987, MNRAS, 227, 1

Kang, X., Jing, Y. P., Mo, H. J., & Börner, G. 2005, ApJ, 631, 21

Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64

Li, C., Jing, Y. P., Kauffmann, G., Börner, G., Kang, X., & Wang, L. 2007, MNRAS, 376, 984

Li, C., Jing, Y. P., Kauffmann, G., Börner, G., White, S. D. M., & Cheng, F. Z. 2006a, MNRAS, 368, 37

Li, C., Kauffmann, G., Jing, Y. P., White, S. D. M., Börner, G., & Cheng, F. Z. 2006b, MNRAS, 368, 21

Lin, W. P., Jing, Y. P., Mao, S., Gao, L., & McCarthy, I. G. 2006, ApJ, 651, 636

Lupton, R. 1993, Statistics in theory and practice (Princeton, N.J., Princeton University Press)

McCracken, H. J., Shanks, T., Metcalfe, N., Fong, R., & Campos, A. 2000, MNRAS, 318, 913

Mobasher, B., Ellis, R. S., & Sharples, R. M. 1986, MNRAS, 223, 11

Norberg, P., Baugh, C. M., Hawkins, E., Maddox, S., et al. 2001, MNRAS, 328, 64

Peebles, P. J. E. 1980, The large-scale structure of the universe (Princeton, N.J., Princeton University Press)

Peebles, P. J. E. 1993, Principles of physical cosmology (Princeton, N.J., Princeton University Press)

Phillipps, S., Fong, R., Fall, R. S. E. S. M., & MacGillivray, H. T. 1978, MNRAS, 182, 673

Roche, N. D., Dunlop, J., & Almaini, O. 2003, MNRAS, 346, 803

Sandage, A., Tammann, G. A., & Yahil, A. 1979, ApJ, 232, 352

Saunders, W., Rowan-Robinson, M., & Lawrence, A. 1992, MNRAS, 258, 134

Schechter, P. 1976, ApJ, 203, 297

Scoccimarro, R. 2004, Phys. Rev. D, 70, 083007

Slosar, A., Seljak, U., & Tasitsiomi, A. 2006, MNRAS, 366, 1455

Springel, V. 2005, MNRAS, 364, 1105

Tegmark, M., et al. 2004, ApJ, 606, 702

Waddington, I., et al. 2007, MNRAS, 381, 1437

Weinberg, D. H., Davé, R., Katz, N., & Hernquist, L. 2004, ApJ, 601, 1

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

Yoshikawa, K., Jing, Y. P., & Börner, G. 2003, ApJ, 590, 654

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

Zehavi, I., et al. 2005, ApJ, 630, 1

Zhao, D., Jing, Y. P., & Börner, G. 2002, ApJ, 581, 876