THE ASTROPHYSICAL JOURNAL, 503:L9–L13, 1998 August 10

ACCURATE FITTING FORMULA FOR THE TWO-POINT CORRELATION FUNCTION OF DARK MATTER HALOS

Y. P. JING
Research Center for the Early Universe, School of Science, University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan; jing@utaphp2.phys.s.u-tokyo.ac.jp
Received 1998 May 15; accepted 1998 June 10; published 1998 July 17

ABSTRACT

An accurate fitting formula is reported for the two-point correlation function $\xi_{hh}(r; M)$ of dark matter halos in hierarchical clustering models. It is valid for the linearly clustering regime, and its accuracy is about 10% in $\xi_{hh}(r; M)$ for the halos with mass $M > (10^{-2}–10^{-3})M_*$, where M_* is the characteristic nonlinear mass. The result is found on the basis of a careful analysis for a large set of scale-free simulations with 256^3 particles. The fitting formula has a weak explicit dependence on the index n of the initial power spectrum but can be equally well applied to the cold dark matter (CDM) cosmological models if the effective index $n_{\rm eff}$ of the CDM power spectrum at the scale of the halo mass replaces the index n. The formula agrees with the analytical formula of Mo & White for massive halos with $M > M_*$, but the Mo & White formula significantly underpredicts $\xi_{hh}(r; M)$ for the less massive halos. The difference between the fitting and the analytical formulae amounts to a factor $\gtrsim 2$ in $\xi_{hh}(r; M)$ for $M = 0.01M_*$. One of the most interesting applications of this fitting formula would be the clustering of galaxies, since the majority of halos hosting galaxies satisfies $M \ll M_*$.

Subject headings: cosmology: theory — dark matter — galaxies: formation — large-scale structure of universe

1. INTRODUCTION

It is generally believed that galaxies are formed within the deep potential wells of virialized dark matter (DM) halos and that clusters of galaxies are recently collapsed objects. The study of the physical properties of DM halos in cosmological models therefore provides important clues to our understanding of the universe. In this Letter, we report a fitting formula for the two-point correlation function $\xi_{hh}(r)$ of the DM halos in hierarchical clustering models. The accuracy of the fit is about 10% in $\xi_{hh}(r; M)$ for a wide range of halo masses.

The two-point correlation function of DM halos has been the subject of many recent attempts at analytical modeling (e.g., Kashlinsky 1987, 1991; Cole & Kaiser 1989; Mann, Heavens, & Peacock 1993; Mo & White 1996, hereafter MW96; Catelan et al. 1998; Porciani et al. 1998), as well as of *N*-body simulation studies (e.g., White et al. 1987; Efstathiou et al. 1988; Bahcall & Cen 1992; Jing et al. 1993; Watanabe, Matsubara, & Suto 1994; Gelb & Bertschinger 1994; Jing, Börner, & Valdarnini 1995; MW96; Mo, Jing, & White 1996). In particular, using the extended Press-Schechter (PS) formalism to calculate the correlation function of DM halos in Lagrangian space (cf. Cole & Kaiser 1989) and mapping from Lagrangian space to Eulerian space within the context of the spherical collapse model, MW96 have derived an analytical expression for $\xi_{hh}(r; M)$:

$$\xi_{hh}(r; M) = b^2(M)\xi_{mm}(r). \tag{1}$$

This should hold in the linearly clustering regime where the mass two-point correlation function $\xi_{mm}(r)$ is less than unity. The bias parameter $b(M)$ is

$$b(M) = 1 + \frac{\nu^2 - 1}{\delta_c} = 1 + \frac{\delta_c}{\sigma^2(M)} - \frac{1}{\delta_c}, \tag{2}$$

where $\sigma(M)$ is the linearly evolved rms density fluctuation of top-hat spheres containing on average a mass M, $\nu \equiv \delta_c/\sigma(M)$, and $\delta_c = 1.68$ (see MW96 and references therein for more details about these quantities). The parameter ν will be called the peak height for convenience. Equations (1) and (2) were found to be in good agreement with their *N*-body results by MW96 and by Mo et al. (1996), but their tests were limited to high-mass halos with $\nu \gtrsim 1$ because of limited mass and force resolutions. The formula has been widely used: from modeling the correlation function of different types of galaxies (e.g., Kauffmann, Nusser, & Steinmetz 1997; Baugh et al. 1998) to interpreting the observed clustering of various extragalactic objects (e.g., Mo et al. 1996; Mo & Fukugita 1996; Matarrese et al. 1997; Steidel et al. 1998; Coles et al. 1998; Fang & Jing 1998).

We have measured the two-point correlation functions for DM halos in a large set of high-resolution *N*-body simulations. Each simulation uses 256^3 particles, and a wide range of hierarchical models are covered: four scale-free models and three representative CDM models. Moreover, each model is simulated with three to four different realizations, and two different box sizes are used for each CDM model. With these simulations of very high accuracy, we can determine $\xi_{hh}(r, M)$ for a wide range of the halo mass M. As a result, we find that the linear bias (eq. [1]) is a good approximation in the linearly clustering regime, but the bias parameter given by equation (2) agrees with the *N*-body results only for massive halos with mass $M \gtrsim M_*$, where M_* is a characteristic nonlinear mass scale defined by $\nu(M_*) = 1$. For less massive halos, the *N*-body results imply significantly higher bias than the analytical prediction equation (2), and the difference in the correlation amplitude amounts to a factor $\gtrsim 2$ for $M = 0.01M_*$. Fortunately, the difference between the *N*-body results and the MW96 formula can be modeled by a simple fitting formula (§ 3). This formula can fit the simulation bias parameter for halo mass $M/M_* \gtrsim 0.01$ with an accuracy of about 5%.

The new findings have profound implications for the formation of large-scale structures. One of the most interesting applications of the fitting formula would be the clustering of galaxies, since local late-type and dwarf galaxies are believed to have mass $\sim 10^{11}\ M_\odot$ and to have formed recently (redshift $z \lesssim 1$; Mo, Mao, & White 1998), while $M_* \sim 10^{13}\ M_\odot$ is expected for the present universe (cf. § 3).

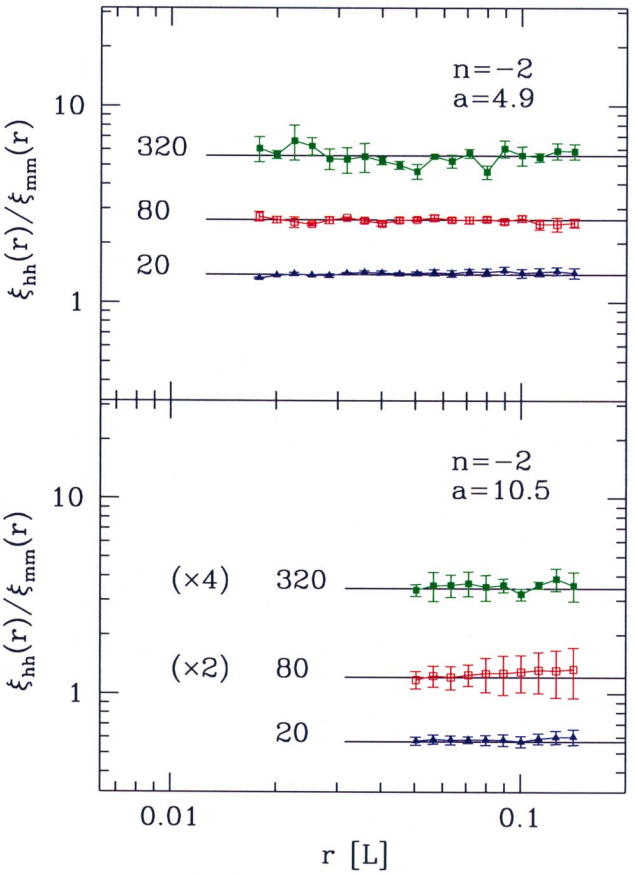

FIG. 1.—The ratio, as a function of the separation r, of the two-point correlation function of the halos to that of the dark matter at two time outputs of the scale-free model with $n = -2$. The separation is in units of the simulation box size L, and only the linear clustering regime is considered. The mass of the halos M, in units of the particle mass, is in the ranges $20 \le M < 40$ (*blue triangles*), $80 \le M < 160$ (*red open squares*), and $320 \le M < 640$ (*green filled squares*), respectively. Because the ratios of the three different masses at the later output (*lower panel*) are very close, for clarity, the ratios for the two larger masses have been multiplied by the factors indicated in parentheses. The solid lines are the mean ratio averaged for different scales.

The simulations will be described in § 2, with emphasis on the aspects relevant to this Letter. In § 3, we will present the correlation function of the halos and the fitting formula. The implications for theories and observations are discussed in § 4.

2. MODELS AND SIMULATIONS

The two-point correlation functions of halos are studied both for scale-free models and for representative CDM models of hierarchical clustering. In the scale-free models, a power-law power spectrum $P(k) \propto k^n$ is used for the initial density fluctuation and the universe is assumed to be Einstein–de Sitter, $\Omega = 1$. Four models with $n = -0.5$, -1.0, -1.5, and -2.0 are studied. Because these models are conceptually simple and exhibit interesting scaling properties, it is relatively easy to understand how physical properties depend on the shape of the power spectrum and, perhaps more importantly, to distinguish the physical effects from numerical artifacts, since the latter should not in general obey the scaling relations. For this reason, we will extensively use the scaling property that the bias parameter b depends only on the halo mass M scaled by M_*, i.e., M/M_* for each n. This scaling property manifests itself when the bias parameter for M/M_* is plotted for different evolution times. Each of our simulations for $n \ge -1.5$ is evolved for 1000 time steps with a total of seven outputs, and that for $n = -2.0$ is evolved for 1362 steps with eight outputs. The output time interval is chosen so that M_* at each successive output is increased by a factor 2.5, and the M_* values (in units of the particle mass) at the first output are 74, 59, 35, and 13 for $n = -0.5$, -1.0, -1.5, and -2.0, respectively. Note that fixing the M_* values is equivalent to fixing the normalization for the power spectra. The time step and the integration variables are taken similarly to Efstathiou et al. (1988). In this Letter we will rely on these scale-free models to understand how the halo-halo correlation depends on the shape of the power spectrum. Then we will examine whether the result obtained from the scale-free models can be applied to CDM models, since CDM models, at least variants thereof, are believed to be close to reality.

Three CDM models are very typical: one is the (ever) standard CDM model (SCDM), one is an open model with $\Omega_0 = 0.3$ and with a vanishing cosmological constant λ_0 (OCDM), and the other is a flat lower density model with $\Omega_0 = 0.3$ and $\lambda_0 = 0.7$ (LCDM). These CDM models are completely fixed with regard to the DM clustering if the initial density power spectrum is fixed. For our simulations, the linear CDM power spectrum of Bardeen et al. (1986) for the primordial Harrison-Zeldovich spectrum is used for the initial condition, which is fixed by the shape parameter $\Gamma = \Omega_0 h$ and the amplitude σ_8 (the rms top-hat density fluctuation at radius 8 h^{-1} Mpc). The values of (Γ, σ_8) are (0.5, 0.62) for SCDM, (0.25, 1) for OCDM, and (0.20, 1) for LCDM.

Each model is simulated with our vectorized P^{3}M code on the Fujitsu VPP300/16R supercomputer at the National Astronomical Observatory of Japan. Each simulation is performed with 256^3 (≈17 million) particles and with good force resolution $\eta \approx 1/2000L$ (where L is the simulation box size). To properly understand the effect of the cosmic variance, three to four independent realizations are generated for each simulation of one model. Furthermore, two different box sizes, 100 h^{-1} Mpc and 300 h^{-1} Mpc, are adopted for each CDM model.

Further details about the code and the simulations will be given in a forthcoming paper (Jing 1998), where many clustering statistics of the dark matter will also be presented. The CDM simulations with box size 100 h^{-1} Mpc were used by Jing & Suto (1998) to study the constraints on cosmological models of the high concentration of the Lyman break galaxies at redshift $z \approx 3$ discovered by Steidel et al. (1998).

3. THE CORRELATION FUNCTION OF THE HALOS AND THE BIAS PARAMETER

The DM halos are identified with the friends-of-friends (FOF) algorithm with a linking parameter 0.2 times the mean particle separation. The halos with at least 20 members are used for the clustering analysis. It is known that the mass defined by the members of such FOF groups is very close to that defined by the spherical overdensity (SO) virialization (Cole & Lacey 1994) and that the mass function of such FOF groups follows the predictions of the PS formalism (e.g., Lacey & Cole 1994; Mo et al. 1996). More importantly, the correlation function of DM halos is quite robust to reasonable halo identification methods, since, for example, the correlation function of the FOF groups is statistically indistinguishable from that of the SO groups (Mo et al. 1996). Therefore, it would suffice to use the FOF groups for the present purpose.

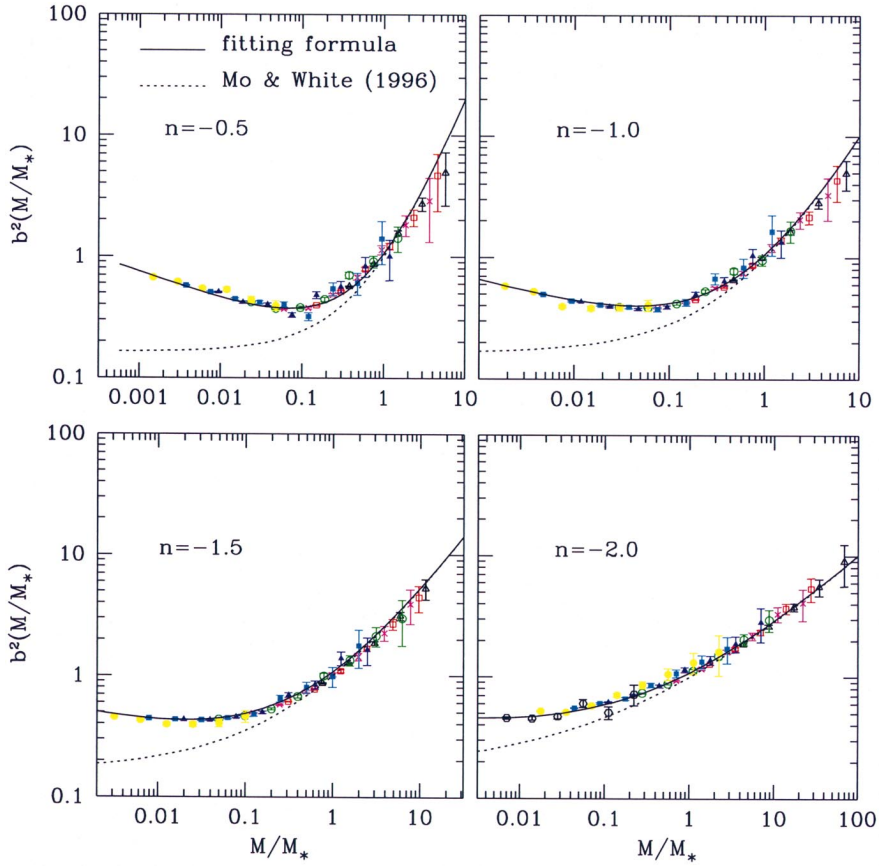

FIG. 2.—The square of the bias parameter as a function of the scaled mass. The results at seven different evolutionary stages are plotted with differently colored symbols. From the early to late outputs, the symbols are, respectively, black open triangles, red open squares, magenta crosses, green open circles, blue solid triangles, cyan solid squares, and yellow solid circles. For $n = -2$, the result for a further output (more clustered) at time step 1362 is added with black hexagons. It is interesting to note that the results from different outputs agree remarkably well. The dotted lines are the prediction of Mo & White (1996), which is in good agreement with the simulation results for $M/M_* \gtrsim 1$, while it significantly underpredicts for $M/M_* \ll 1$. The solid lines are from the simple formula found in this Letter (eq. [3]), which can accurately fit the simulation results.

The first concern of this work is whether the bias of the halos is linear (eq. [1]) in the linearly clustering regime. We have calculated the ratio of $\xi_{hh}(r)$ to $\xi_{mm}(r)$ for every simulation output. Some examples that are also typical are shown in Figure 1, which shows the ratio at two different outputs of the $n = -2$ simulation for three different halo masses. The results are plotted only for the linear clustering regime, i.e., $\xi_{mm}(r) < 1$, as only this regime is considered here. Error bars are calculated by averaging over the different realizations. It is remarkable that the ratio is a constant within the 1 σ error bars; i.e., the bias is linear and the bias parameter b is a function of M only. This statement is consistent with many previous studies but is shown here with higher accuracy.

Since $b(M)$ depends only on M, we predict that for a scale-free model, if the mass M is scaled by the characteristic mass M_*, the bias parameter b depends only on the scaled mass M/M_*. Interestingly, by the definitions of ν and M_*, the scaled mass obeys a simple relation to the peak height ν, i.e., $\nu = (M/M_*)^{n+3/6}$. Thus, it is also very convenient to compare the simulation data of $b(M/M_*)$ with the formula of MW96 (cf. eq. [2]). In Figure 2, we plot the square of the bias parameter b^2 as a function of the scaled mass M/M_*. The results for all simulation outputs are depicted. This is equivalent to a factor $\sim 10^4$ in the halo mass resolution. It is very remarkable that the $b^2(M/M_*)$ results agree very well for different evolutionary outputs. The precise scaling behavior given by the simulation data assures that any numerical artifacts have negligible effect on the results of Figure 2.

Now we compare our simulation results with the prediction of MW96. The MW96 predictions are drawn in Figure 2 as dotted lines. It is interesting to see that the MW96 formula agrees well (within ~ 1 σ uncertainty) with the simulation results for massive halos, i.e., $M/M_* > 1$. However, at $M/M_* \approx 1$, the simulation data start to deviate from the MW96 prediction, and the deviation increases with the decrease of the scaled mass (or, equivalently, with the decrease of the peak height). At $M/M_* = 0.01$, the simulation result is about 2–4 times higher than the prediction. Considering the important role played by the correlation functions of the halos in the cosmological studies, we have searched for a fitting formula for $b(M/M_*)$. The formula

$$b(M) = \left(\frac{0.5}{\nu^4} + 1\right)^{(0.06-0.02n)}\left(1 + \frac{\nu^2 - 1}{\delta_c}\right), \quad \nu = (M/M_*)^{n+3/6}, \tag{3}$$

can fit the simulation results (Fig. 2) with an accuracy of about 5% for the halo mass that the simulations can probe; i.e., $M/M_* \gtrsim 10^{-3}$ for $n = -0.5$, $M/M_* \gtrsim 2 \times 10^{-3}$ for $n = -1.0$, $M/M_* \gtrsim 3 \times 10^{-3}$ for $n = -1.5$, and $M/M_* \gtrsim 10^{-2}$ for $n = -2.0$. The fitting formula recovers the MW96 analytical for-

mula (eq. [2]) for high-mass halos at $\nu \gtrsim 1$. The deviation of the MW96 prediction from the simulation results at small mass is accounted for by the factor $[(0.5/\nu^4) + 1]^{(0.06-0.02n)}$.

Since the scale-free models are only approximations to the real universe at some specific scales, it is very important to consider more realistic models. In Figure 3, we present the $b^2(M)$ for the halos in the three CDM models. The blue squares are determined from the simulations of box size 300 h^{-1} Mpc, and the red triangles are from those of the smaller boxes. Consistent with the scale-free models, the simulation data are significantly higher than the MW96 prediction (*dotted lines*) for masses less than M_*, where M_* is about $10^{13}\ h^{-1}\ M_\odot$ for the SCDM and about $2 \times 10^{13}\ h^{-1}\ M_\odot$ for the low-density models. Since the slopes of the CDM power spectra change with scale, the fitting formula (eq. [3]) is not directly applicable. Fortunately, the fitting formula depends very weakly on the power index n, and it can describe the CDM data very accurately if an effective index $n_{\rm eff}$ replaces n in equation (3) and the original definition of $\nu = \delta_c/\sigma(M)$ is used for ν. The effective index is defined as the slope of $P(k)$ at the halo mass M:

$$n_{\rm eff} = \frac{d \ln P(k)}{d \ln k}\Bigg|_{k=2\pi/R}, \quad R = \left(\frac{3M}{4\pi\bar{\rho}}\right)^{1/3}, \qquad (4)$$

where $\bar{\rho}$ is the mean density of the universe. The solid lines in Figure 3 are predicted in this way. They agree very well with the simulation results. The fact that $n_{\rm eff}$ changes very slowly with the mass M for $M \leq M_*$ might be the main reason why the fitting formula can work very well for CDM models after the simple modification.

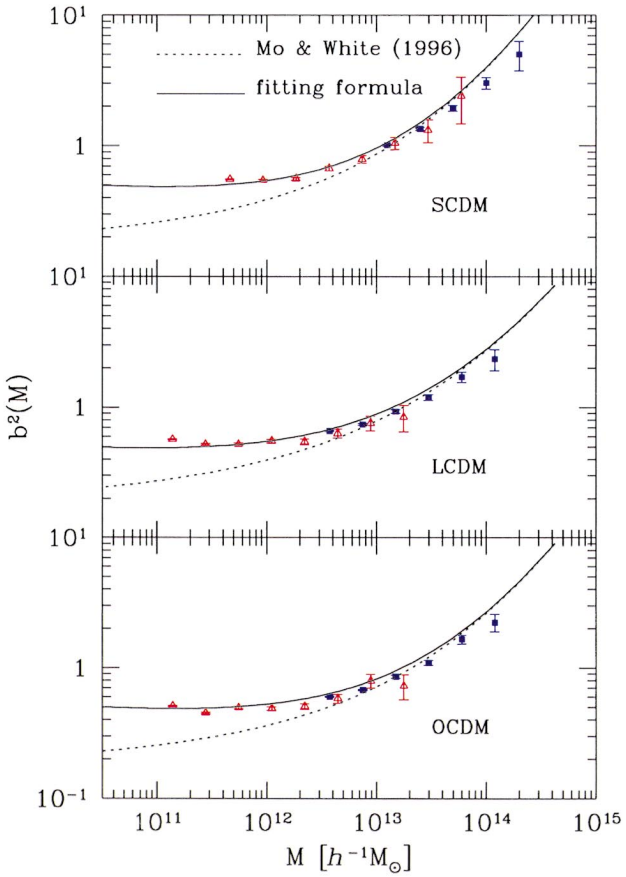

FIG. 3.—The square of the bias parameter as a function of the halo mass for three typical CDM models. The blue squares are for the simulations of the larger box, and the red triangles are for the smaller box. The dotted lines are the prediction of Mo & White (1996), which showed a similar behavior to that found in the scale-free models when compared with the simulation data. The solid lines, which agree quite well with the simulation data, are the prediction of the fitting formula in this Letter with the index n in eq. (3) replaced with the effective one $n_{\rm eff}$ at the halo mass scale (see text).

4. DISCUSSION AND CONCLUSIONS

We have determined the correlation function of DM halos for the hitherto largest set of high-resolution N-body simulations. The excellent scaling exhibited in the bias parameter $b(M/M_*)$ at the different evolutionary outputs for the scale-free models assures that the results in Figures 2 and 3 and are physical and are not contaminated by numerical artifacts. The simulation results are in good agreement with the formula of Mo & White (1996) for massive halos $M \gtrsim M_*$. However, for less massive halos, the simulation results are significantly higher. The MW96 formula was found to be in good agreement with the results of 100^3 particle scale-free simulations by MW96 and with the results of 128^3 particle CDM simulations of box size ~300 h^{-1} Mpc by Mo et al. (1996). However, their tests were limited to halos with $M \gtrsim M_*$ because of the relatively poorer mass resolutions. The results found here therefore do not contradict with, but in fact support and extend, the previous N-body tests.

The fitting formula (eq. [3]) we found for $b(M)$ is accurate for halo masses $M/M_* > 10^{-2}$–10^{-3} with only about 5% error. The formula can be applied both to the scale-free models and to the CDM models. In the latter case, the index n in equation (3) should be replaced with the effective power spectrum index $n_{\rm eff}$ (eq. [4]) and the original definition of $\nu = \delta_c/\sigma(M)$ is used for ν. This fitting formula could have many important applications for the studies of large-scale structures. One of them would be to predict and to interpret the clustering of late-type galaxies and dwarf galaxies both in real observations and in analytical modeling of galaxy formation, since these galaxies are believed to have formed recently ($z \lesssim 1$; cf. Mo et al. 1998) with halo masses much less than M_*.

At present, we do not know the *exact* reasons that cause the MW96 formula to fail at the small halo masses $M/M_* \ll 1$. In the derivation of MW96, two main assumptions are that (1) the halo formation is determined by the peak height through the extended PS formalism and that (2) the mapping of the halo clustering pattern from Lagrangian space to Eulerian space is local and linear (with the spherical collapse model). The first assumption could break down more seriously for low peak height halos, because the tidal force plays a more important role in their formation; although, as Katz, Quinn, & Gelb (1993; see also Katz et al. 1994) pointed out, the peak height is not the sole parameter even for the formation of high-peak halos. The validity of the local linear mapping was recently questioned by Catelan, Matarrese, & Porciani (1998) in a different context. Unfortunately, their result is not directly applicable to this Letter. In relation to the gentle rise observed for the bias parameter at the small M/M_* (Fig. 2), we have visually inspected the spatial distribution for halos with $M/M_* \approx 3 \times 10^{-3}$ ($b \approx 0.8$), $M/M_* \approx 10^{-1}$ ($b \approx 0.6$), and $M/M_* \approx 1$ ($b \approx 1$) in one late output of the $n = -0.5$ model. The small and large halos appear to delineate filamentary structures more closely than the median-mass halos, consistent with the measured b. This however

might hint that either or both of the two assumptions are violated for the small halos, since the small halos are otherwise expected to be more preferentially located in low-density regions. Obviously, it would be very interesting to find out why the MW96 formula fails. We will examine this question more closely in a future paper.

I am indebted to Gerhard Börner and Yasushi Suto for their critical reading of earlier versions of the manuscript. Constructive comments by them, by Houjun Mo, and by an anonymous referee improved the presentation of this Letter. Yasushi Suto is thanked for his invaluable assistance and continuing encouragement; it would be impossible to get these very large, high-quality simulations without his help. It is also my pleasure to acknowledge the JSPS foundation for a postdoctoral fellowship. The simulations were carried out on VPP/16R and VX/4R at the Astronomical Data Analysis Center of the National Astronomical Observatory, Japan.

REFERENCES

Bahcall, N. A., & Cen, R. Y. 1992, ApJ, 398, L81
Bardeen, J., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Baugh, C. M., Cole, S., Frenk, C. S., & Lacey, C. G. 1998, ApJ, 498, 504
Catelan, P., Lucchin, F., Matarrese, S., & Porciani, C. 1998, MNRAS, in press (astro-ph/9708067)
Catelan, P., Matarrese, S., & Porciani, C. 1998, ApJ, 502, L1
Gelb, J. M., & Bertschinger, E. 1994, ApJ, 436, 491
Cole, S., & Kaiser, N. 1989, MNRAS, 237, 1127
Cole, S., & Lacey, C. 1996, MNRAS, 281, 716
Coles, P., Lucchin, F., Matarrese, S., & Moscardini, L. 1998, MNRAS, submitted (astro-ph/9803197)
Efstathiou, G., Frenk, C. S., White, S. D. M., & Davis, M. 1988, MNRAS, 235, 715
Fang, L. Z., & Jing, Y. P. 1998, ApJ, 502, L95
Jing, Y. P. 1998, in preparation
Jing, Y. P., Börner, G., & Valdarnini, R. 1995, MNRAS, 277, 630
Jing, Y. P., Mo, H. J., Börner, G., & Fang, L. Z. 1993, ApJ, 411, 450
Jing, Y. P., & Suto, Y. 1998, ApJ, 494, L5
Kashlinsky, A. 1987, ApJ, 317, 19
Kashlinsky, A. 1991, ApJ, 376, L5
Katz, N., Quinn, T., Bertschinger, E., & Gelb, J. M. 1994, MNRAS, 270, L71
Katz, N., Quinn, T., & Gelb, J. M. 1993, MNRAS, 265, 689
Kauffmann, G., Nusser, A., & Steinmetz, M. 1997, MNRAS, 286, 795
Lacey, C., & Cole, S. 1994, MNRAS, 271, 676
Mann, R. G., Heavens, A. F., & Peacock, J. A. 1993, MNRAS, 263, 789
Matarrese, S., Coles, P., Lucchin, F., & Moscardini, L. 1997, MNRAS, 286, 115
Mo, H. J., & Fukugita, M. 1996, ApJ, 467, L9
Mo, H. J., Jing, Y. P., & White, S. D. M. 1996, MNRAS, 282,1096
Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 296, 319
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347 (MW96)
Porciani, C., Matarrese, S., Lucchin, F., & Catelan, P. 1998, MNRAS, in press (astro-ph/9801290)
Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., Pettini, M., & Kellogg, M. 1998, ApJ, 492, 428
Watanabe, T., Matsubara, T., & Suto, Y. 1994, ApJ, 432, 17
White, S. D. M., Frenk, C. S., Davis, M., & Efstathiou, G. 1987, ApJ, 313, 505