THE ASTROPHYSICAL JOURNAL, 538:83–91, 2000 July 20

PHYSICAL BIAS OF GALAXIES FROM LARGE-SCALE HYDRODYNAMIC SIMULATIONS

RENYUE CEN[1] AND JEREMIAH P. OSTRIKER[2]
Received 1998 September 29; accepted 2000 February 29

ABSTRACT

We analyze a new large-scale (100 h^{-1} Mpc) numerical hydrodynamic simulation of the popular ΛCDM cosmological model, including in our treatment dark matter, gas, and star formation, on the basis of standard physical processes. The method, applied with a numerical resolution of $<200\ h^{-1}$ kpc (which is still quite coarse for following individual galaxies, especially in dense regions), attempts to estimate where and when galaxies form. We then compare the smoothed galaxy distribution with the smoothed mass distribution to determine the "bias," defined as $b \equiv (\delta M/M)_{\rm gal}/(\delta M/M)_{\rm total}$, on scales that are large compared to the code numerical resolution (on the basis of resolution tests given in the Appendix of this paper). We find that (holding all variables constant except the quoted one) bias increases with decreasing scale, with increasing galactic age or metallicity, and with increasing redshift of observations. At the 8 h^{-1} Mpc fiducial comoving scale, bias (for bright regions) is 1.35 at $z = 0$, reaching to 3.6 at $z = 3$, both numbers being consistent with extant observations. We also find that (10–20) h^{-1} Mpc voids in the distribution of luminous objects are as observed (i.e., observed voids are not an argument against cold dark matter [CDM]–like models), and finally that the younger systems should show a colder Hubble flow than do the early-type galaxies (a testable proposition). Surprisingly, little evolution is found in the amplitude of the smoothed galaxy-galaxy correlation function (as a function of *comoving* separation). Testing this prediction against observations will allow a comparison between this work and that of Kauffmann et al., which is based on a different physical modeling method.

Subject headings: galaxies: clusters: general — galaxies: formation — large-scale structure of universe — methods: *n*-body simulations

1. INTRODUCTION

The spatial distribution of galaxies provides the core data on which cosmological theories for the growth of structure are based. Spatial correlations, peculiar velocities, voids, etc. have provided the primary evidence for the growth of structure due to gravitational instabilities. However, the theories developed to model this growth of structure largely treat the collisionless dark matter, thought to provide the bulk of the mass density in the universe, rather than the observables, the stellar parts of galaxies. Current estimates would put the ratio of the mass densities of these two components at about $\Omega_{\rm dm}/\Omega_* \approx 0.3/0.003 \approx 100$, so the potential fluctuations (except on the smallest scales, $\Delta r < 10^4$ pc) are dominated by dark matter mass fluctuations. To accommodate this dichotomy between observables and computables, the concept of "bias," b, was developed (e.g., Davis et al. 1985) to bridge the gap. Treated at first as (what it was) a convenient way of parameterizing our ignorance, it has developed a life of its own.

A common definition would be based on the relation between the number density of galaxies, N_g (with, e.g., $|m_V| > |m_{V,0}|$), or the total mass density in galaxies, ρ_g, and the total mass in a given region, $\rho_{\rm tot}$, with bias b_l in regions smoothed by a top-hat smoothing length l, defined by

$$\left\langle \left(\frac{\delta N_g}{N_g}\right)^2 \right\rangle_l \approx \left\langle \left(\frac{\delta \rho_g}{\rho_g}\right)^2 \right\rangle_l \equiv b_l^2 \left\langle \left(\frac{\delta \rho_{\rm tot}}{\rho_{\rm tot}}\right)^2 \right\rangle_l . \quad (1)$$

Here we have explicitly put in the scale dependence, although in primitive treatments bias is sometimes treated as a number. The realization that b_l is both stochastic and scale dependent has grown recently (for discussions of analytic theories of biasing, see Scherrer & Weinberg 1998 and Pen 1998), but, of course, this must be the case. Blanton et al. (1999) have studied our simulations, taking a broader perspective and allowing, on the right-hand side of equation (1), a dependence on variables other than mass density. We restrict ourselves in the current exercise to the conventional parameterization, but we examine how b, defined by second equality in equation (1), may depend on spatial scale, galaxy age, and other galactic properties. Our simulations do not have sufficient spatial resolution to identify individual galaxies (except in low-density regions, where they are adequate), due to a tendency of the present simulation to merge distinct systems in high-density regions. However, galaxy mass is conserved during mergers, so the smoothed galaxy density (smoothing scale 500 h^{-1} kpc) is the variable we compare with the similarly smoothed dark matter density. Since this smoothing scale is small compared to the correlation lengths we examine ($>1\ h^{-1}$ Mpc), we believe that our numerical resolution is adequate (but far from ideal) for the questions examined. As a test, given in Appendix A, a comparison of two simulations with length scale resolutions differing by a factor of 2 reassuringly show no difference in bias on the scales studied, within the statistical errors ($\sim 10\%$).

Nature has determined bias, not by providing us with a mathematically convenient functional form, but through the physical process of galaxy formation. By most current accounts, the bulk of the baryons have not condensed into stellar systems, at present having a ratio of gas to stars $\Omega_b/\Omega_*(\approx 0.04/0.003) \geq 10$. Thus, galaxies appear to have formed only under favorable conditions. Both simple physi-

[1] Princeton University Observatory, Princeton University, Princeton, NJ 08544; cen@astro.princeton.edu.
[2] Princeton University Observatory, Princeton University, Princeton, NJ 08544; jpo@astro.princeton.edu.

cal theory and observations indicate greater formation efficiencies in regions of higher density, where cooling processes were more efficient. In the Galaxy, the observed "Schmidt law" (Schmidt 1959; see also Kennicutt 1989 for an update), stating that star formation rates scale roughly as the locally averaged value of $\rho_{\rm gas}^2$, still seems to have some approximate empirical validity. It is based, of course, on the physical fact that for the dominant collisionally excited line or continuum cooling processes, the gas cooling rate per unit volume, thought to be the rate-limiting step, is proportional to $\rho_{\rm gas}^2$.

Cosmological observations would also seem to support such a picture, since the void regions are believed to be underdense (from simulations) by a factor of 3–5 in total mass density as compared to the average, whereas empirically, these regions have a galaxy density far below this (Peebles 1993) in terms of average galaxy density. Conversely, in the great clusters of galaxies, perhaps 10% of the total baryonic mass is in stellar form, indicating above average efficiency of galaxy formation in these very overdense regions.

One reasonable approach to determining bias is to begin by noting directly that most stellar galaxies live within massive halos. Then one can compute the distribution of dark matter halos and estimate which of these will contain which type of stellar systems. Much work has been done following this promising track in recent years either using large N-body simulations or combining semianalytic dark matter treatments with detailed hydrodynamic simulations (Cole et al. 1994a; Kauffmann, Nusser, & Steinmetz 1997). These analyses have produced results that are consistent with many observations.

However, direct numerical simulations can be made on an ab initio basis that combine the physical gasdynamic processes used in the hydrodynamic simulations of Katz, Hernquist, & Weinberg (1992) or Steinmetz (1996) with large-scale numerical simulations of dark matter pioneered by Davis, Efstathiou, Frenk, and White (Davis et al. 1985; Frenk et al. 1985; White et al. 1987). Such calculations could, in principle, determine "bias" by direct computations. We have attempted to do that in earlier work (Cen & Ostriker 1992, 1993) but, as is well known, the difficulties and uncertainties are formidable. These are of two kinds: inadequate physical modeling and insufficient numerical resolution.

In the last five years, we have made quite significant improvements in both our physical modeling and the numerical resolution that we can achieve. First, we have upgraded from the somewhat diffusive aerospace gasdynamics-based Eulerian hydrocode (Cen 1992) to a shock-capturing, high-order, total variation diminishing (TVD) hydrocode (Ryu et al. 1993). A very important additional ingredient in the new TVD code is the implementation of a new, entropy variable into the conventional TVD scheme (Harten 1983). The true spatial resolution for a given mesh has increased by a factor of about 2 (Kang et al. 1994), and unphysical spurious heating is removed, allowing for a much more accurate treatment in lower temperature regions. We have also gained a factor of 2.5 through Moore's law (increasing computational power), so that the overall gain is approximately a factor of 5 from >1 h^{-1} Mpc to 200 h^{-1} kpc. In particular, our current simulations have a nominal spatial resolution that is small compared to the typical distance between galaxies (see Fig. 1), which was not the case earlier; therefore, while we are still totally unable to say anything about the internal structural properties of galaxies, we should be able to specify galaxy integral properties: we feel that the resolution suffices to determine where and when galaxies of normal range will form, and to determine, locally, the mass density in galaxies, while we remain unable to say very much about the number density of galaxies.

Second, we have substantially improved the realism of the input physics, aside from a seemingly better cosmological model. First, we have added more relevant microphysical processes due to elements other than hydrogen and helium. This directly affects cooling/heating, optical opacity, and the radiation field, and indirectly affects them through complex interplays among these physical processes and gravity. In addition, in our galaxy formation algorithm we now allow not only energy feedback from stars but also matter (including metals) ejection into the intergalactic medium (IGM). Finally, we separately but simultaneously follow the dynamics of metals in the IGM produced during the feedback processes. The algorithm for identifying regions of star formation, which has remained the same, has also been used by other researchers to make detailed hydrodynamic simulations: e.g., Katz et al. (1992), Katz, Weinberg, & Hernquist (1996), Steinmetz (1996), Gnedin & Ostriker (1997), and Abel et al. (1998). In a region of sufficient overdensity, if three conditions are *simultaneously* satisfied, we assume that an already existing collapse cannot be reversed and that collapsed objects of some kind (stars, star clusters, or molecular clouds) will form. The criteria in detail are:

$$\begin{aligned} &1.\ \nabla\cdot\boldsymbol{v} < 0\ , \\ &2.\ M_{\rm gas} > M_{\rm Jeans}\ , \\ &3.\ t_{\rm cooling} < t_{\rm dynamical}\ . \end{aligned} \tag{2}$$

If all are estimated to be satisfied within a cell, then a mass in gas equal to $\Delta M_g = -C_* M_g \Delta t/t_{\rm dyn}$ in a time step Δt is removed from the cell (having gas mass M_g) and put into a collisionless particle of mass $\Delta M_* = -M_g$, where C_* is the star formation efficiency, with a value of 0.20. That particle is labeled with the time of formation, the metallicity of the gas from which it was made, and the density of the cell in which is was made. These "stellar" particles have a typical mass of $10^7\ M_\odot$ and can be thought of as clusters within which star formation of essentially coeval stars will occur. They are followed after birth by the same code that tracks the collisionless dark matter particles. We find that (except in high-density regions) these particles group themselves into galaxy-like objects (see Fig. 1), which have a total mass, a mean time of formation (with dispersion), and a mean metallicity (with dispersion). Within each "galaxy" of defined mean age there will be stellar particles having a wide range of ages as the computed galaxies (like our own) are assembled over time. Even in high-density regions, we can smooth over some spatial scale and define spatial fields having "stellar" density, age, and metallicity.

The top panel of Figure 1 shows the dark matter density for a random slice of size $100 \times 100\ h^{-2}$ Mpc^2 with a thickness of three cells ($586\ h^{-1}$ kpc). Each small tick mark of the top panel has a size of $6.25\ h^{-1}$Mpc. The three bottom pairs of panels show three galaxy groups in this slice in an enlarged display; two separate panels are shown for each selected group, with the top panel showing the galaxy

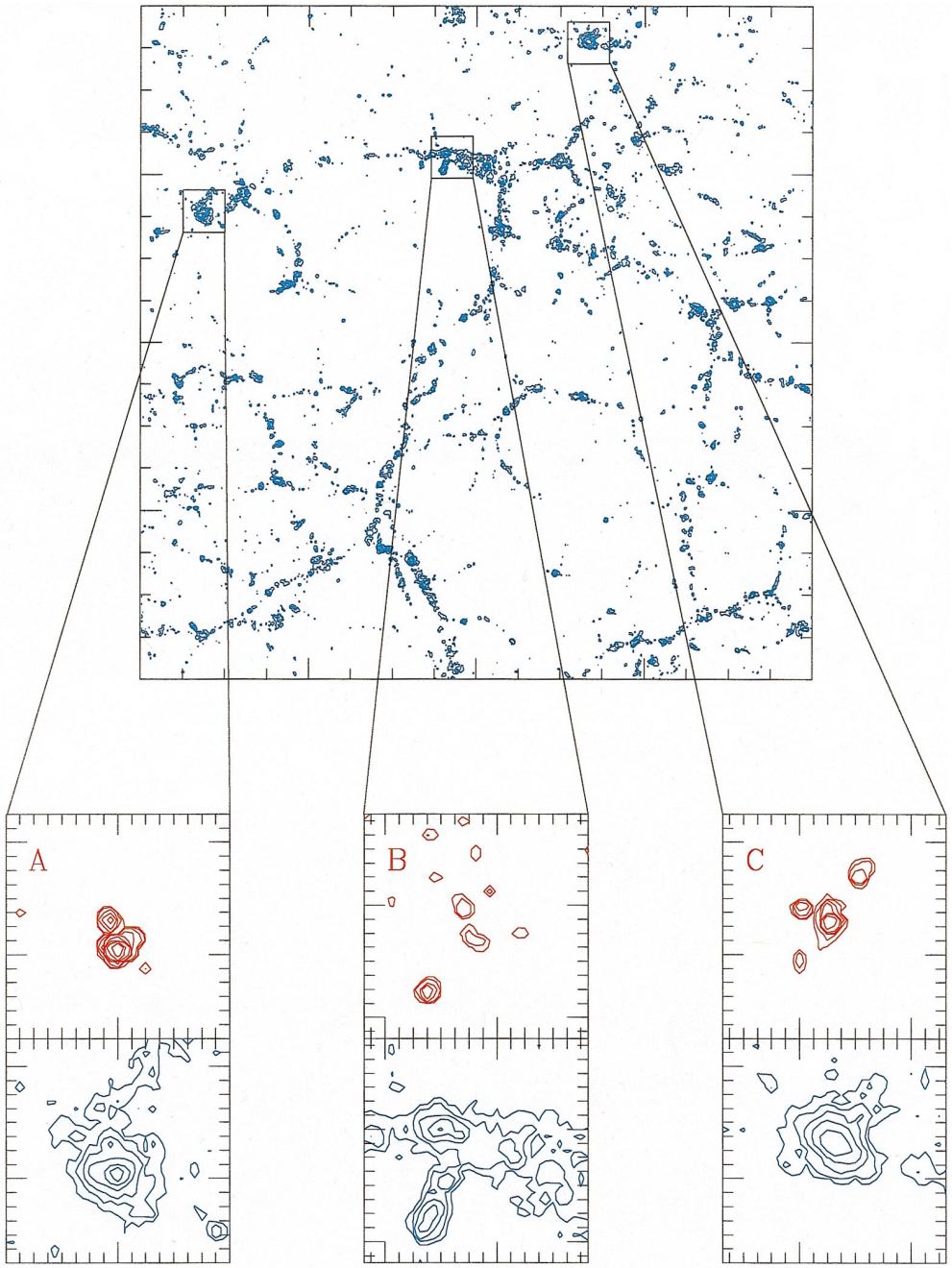

FIG. 1.—Top panel shows dark matter density for a random slice of size $100 \times 100\ h^{-2}$ Mpc^2 and thickness of three cells ($586\ h^{-1}$ kpc). Each small tick mark of the top panel has a size of $6.25\ h^{-1}$ Mpc. The three bottom pairs of panels show three galaxy groups in this slice in an enlarged display; two separate panels are shown for each selected galaxies, the top panel showing the galaxy density contour and the bottom panel showing the dark matter density contour. Each tick mark for the bottom panels has a size of $391\ h^{-1}$ kpc, corresponding to two cells, which is approximately the resolution of the code (Cen & Ostriker 1999).

density contour and the bottom panel showing the dark matter density contour. Each small panel has a size of $6.25 \times 6.25\ h^{-2}$ Mpc^2, and each tick mark for the bottom panels has a size of $391\ h^{-1}$ kpc, corresponding to two cells, which is approximately the resolution of the code (Cen & Ostriker 1999). For group A, we see that the galaxy particles have grouped themselves into two quite separated galactic lumps, whereas the dark matter contour shows only one major lump. For group B, several galactic lumps with varying mass are seen, which roughly follow the distribution of the dark matter, but the difference between galaxy density and dark matter is noticeable. In group C the two (out of four) galactic lumps in the middle are clearly separated, while they are embedded in a common halo. Clearly, any galaxy pairs that are separated by $1\ h^{-1}$ Mpc are resolved by the code. We note the fact that we often see distinct galactic-mass objects swimming separately in a common halo. This is not due to differences in numerical resolution between dark matter and galaxy particles: the galaxy particles and dark matter particles are followed with an identical PM code. Rather, this occurs because the galaxy particles formed from gaseous accumulations that had undergone radiative cooling and contraction. In the objects shown as distinct lumps, all normal dynamical processes are active, including dynamical friction and mergers (which of course are somewhat overestimated).

Considering the substantial improvements in the code, it seems worthwhile to return to the problem of bias using new simulations. The combination of the improvements already noted allows an identification of individual galaxies in low-density regions. In addition, some new questions can now be addressed. For example, we can now study many problems concerning the distribution of metals in a self-consistent fashion. In this paper, we study the relative distributions of galaxy mass density and total mass density, and their evolution in a cold dark matter model with a cosmological constant (ΛCDM). Some aspects (scale dependence of bias) of this work have been reported on in Blanton et al. (1999). Here we focus on other observables, including dependence on epoch.

The new simulations allow us to address the redshift dependence of this bias, which has also been recently addressed by Katz et al. (1999). We will also examine the dependence on galaxy mass and metallicity and the bias of other constituents, such as uncondensed gas in various temperature ranges. Because of the very considerable cost of these large-scale simulations, we have only completed the analysis of one currently favored cosmological model. Future work will allow us to compare bias among models.

After briefly describing further properties of the computational method and the specific cosmological model considered in § 2, we present the results in § 3. Section 3.1 is devoted to galaxy bias at redshift zero, while § 3.2 examines the evolution of galaxy bias with redshift. Finally, our conclusions are summarized in § 4. Appendix A presents an examination of the effects of the finite numerical resolution on our results by comparing output of two simulations having spatial resolution differing by a factor of 2.

2. METHOD AND COSMOLOGICAL MODEL

We model galaxy formation as described earlier. Self-consistently, feedback into the IGM from young stars in these galaxy particles are allowed, in three related forms: supernova mechanical energy output, UV photon output, and mass ejection from the same supernova explosions.

We adopt a cold dark matter model with a cosmological constant (ΛCDM), close to the concordance model of Ostriker & Steinhardt (1995) and the preferred model of Krauss & Turner (1995), with the following parameters: $\Omega = 0.37$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.63$, $h = 0.70$, and $\sigma_8 = 0.80$. The model is also consistent with that favored by recent Type Ia supernova observations (Schmidt et al. 1998). This model is normalized to both *COBE* and the observed cluster abundance at zero redshift (Eke, Cole, & Frenk 1996). The current age of the universe of this model is 12.7 Gyr, consistent with recent age constraints from the latest globular cluster observations/interpretations (see, e.g., Salaris, Degl'Innocenti, & Weiss 1997). The choice of the Hubble constant is in agreement with current observations; it appears that $H_{0,\rm obs} = 65 \pm 10$ km s^{-1} Mpc^{-1} can account for the distribution of the current data from various measurements (see, e.g., Trimble 1997). The choice of $\Omega_b = 0.024\ h^{-2}$ is consistent with current observations of Tytler, Fan, & Burles (1996).

The box size of the primary simulation is 100 h^{-1} Mpc, and there are 512^3 cells (on a uniform mesh), giving a nominal resolution of 197 h^{-1} kpc. However, the actual resolution, as determined by extensive tests (Cen & Ostriker 1998), is approximately a factor of 1.1–1.7 worse than this, depending on the region in question. There are 256^3 dark matter particles used, the mass of each particles being $5.3 \times 10^9\ h^{-1}\ M_\odot$. The typical galaxy particle has a mass of $10^7\ M_\odot$. In addition to this primary box, which will be used for most of our analysis, we have made a higher resolution run with box size 50 h^{-1} Mpc, with the same number of cells, which will be used for resolution-calibration purposes in Appendix A.

3. RESULTS

3.1. *Bias of Galaxies at Redshift Zero*

Figure 2*a* shows the bias of galaxies for four sets ($4 \times 25\%$ in mass) of galaxies ordered by formation redshift ($z = 0.0$–0.8, 0.8–1.0, 1.0–1.3, and 1.3–10.0), as well as all galaxies at $z = 0$, as a function of top-hat smoothing radius. All scales depicted are considerably larger than our resolution limit; Appendix A indicates that numerical-resolution problems should not be important for the results discussed here. As an important aside, we note that 50% of the galaxy mass was formed between redshifts 0.8 and 1.3, with median galaxy formation epoch $z = 1.0$, consistent with Madau's (1997) analysis of the Hubble Deep Field.

Since all the particles in a given region are tagged with the epoch of formation, the mean (mass-weighted) age of any region is computable at any time. The two youngest sets of galaxies with formation epochs at $z < 1.0$ (Fig. 2*a*, *dotted and short-dashed curves*) are nearly unbiased at large smoothing lengths, e.g., $r_{\rm TH} \sim 15\ h^{-1}$ Mpc (above which the limited box size causes the estimate to be inaccurate) and have a bias of ~ 2.0 at $r_{\rm TH} \sim 1\ h^{-1}$ Mpc, below which our simulation fails. The next set of galaxies that formed at $z = 1.0$–1.3 (*long-dashed curve*), is more strongly biased than the two youngest sets of galaxies. The oldest galaxies (*dot-dashed curve*) are most strongly biased, with a bias of ~ 5.5 at $r_{\rm TH} \sim 1\ h^{-1}$ Mpc, and are still significantly biased at large scales, e.g., ~ 3.0 at $r_{\rm TH} \sim 10\ h^{-1}$ Mpc. If we think of ellipticals as being the stellar systems within which most star formation was completed earliest (cf. Searle & Sargent 1972), they should represent, in our terms, the oldest systems. We know that these are very overabundant in the rare rich clusters, i.e., in regions of very high overdensity, and so are strongly biased even on large scales (Kaiser 1984). We note from Figure 2*a* that at the fiducial smoothing scale of 8 h^{-1} Mpc, the bias for all galaxies is 1.35, given $\sigma_{8,\rm gal} = 1.08$ (for $\sigma_{8,\rm mass} = 0.80$), consistent with the observed value of $\sigma_8 = 1.20 \pm 0.18$ (Loveday et al. 1996). (Note that the observed value of σ_8 is for bright galaxies in the APM survey; this is more appropriate for our comparison because the computed galaxy fluctuation is mass-weighted, thus heavily weighted by the most massive, bright galaxies.)

Figure 2*b* shows directly the autocorrelation functions for the stellar mass density of all galaxies, and the mass and the stellar mass densities of galaxy subsets ordered by formation redshift; the results are consistent with Figure 2*a*. To summarize, all types of galaxies as well as all galaxies are biased over matter on all scales, with two separate trends: (1) at any epoch, older objects are more strongly clustered than younger ones, and (2) bias decreases with increasing scale. Both trends are consistent with our earlier work (Cen & Ostriker 1992). A question has arisen as to whether anti-bias occurs in these simulations, as it does in some quasi-analytic modeling. The answer is "yes," but for different reasons. In dark matter simulations, it is found (Kravtsov &

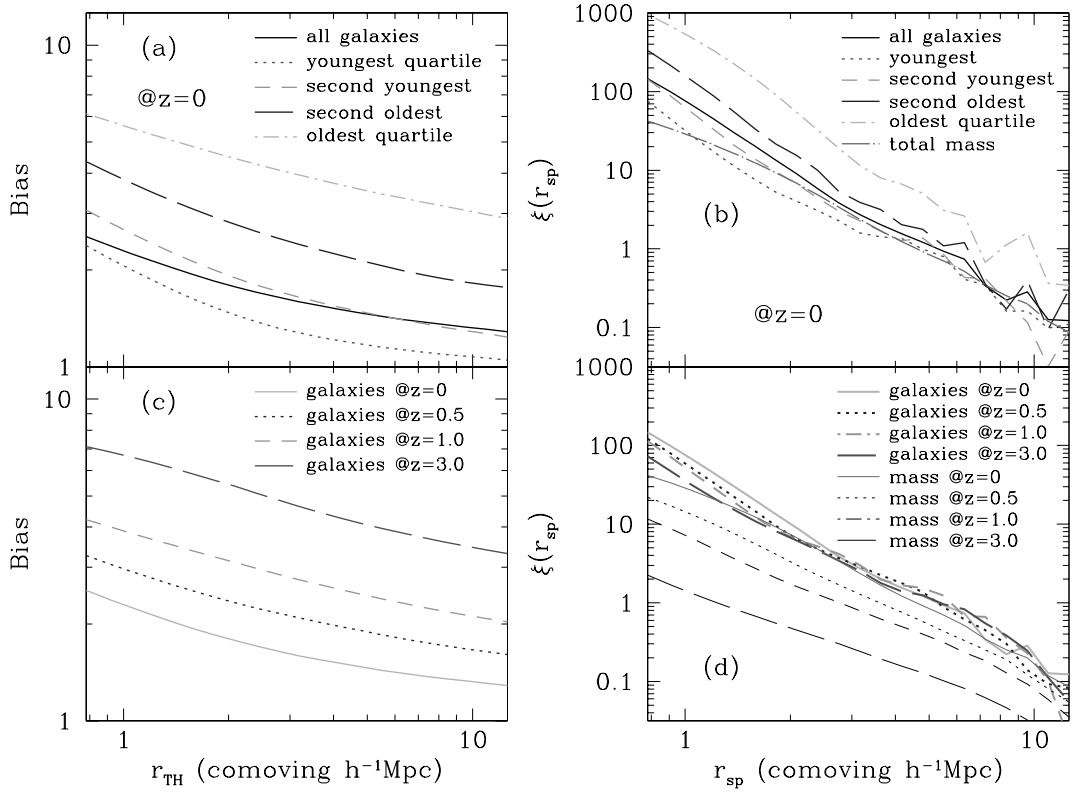

FIG. 2.—(*a*) Bias of galaxies for four sets (4 × 25% in mass) of galaxies ordered by formation time (z = 0.0–0.8, 0.8–1.0, 1.0–1.3, 1.3–10.0), as well as all galaxies at $z = 0$. (*b*) Autocorrelation functions for all galaxies; mass and galaxy subsets are sorted by formation epoch at zero redshift. (*c*) Bias as a function of radius at four different redshifts. (*d*) Autocorrelation functions at various epochs for all galaxies and mass.

Klypin 1999) that in dense regions halos will merge, with the expectation that the embedded galaxies will merge as well, reducing the number density of galaxies (but not the light in galaxies). This effect also occurs in our simulations, but, as is evident from Figure 1, a merger of halos does not necessarily lead to mergers of the dissipational galactic components. We find antibias for young systems in dense regions (Blanton et al. 2000) for the same reason that star-forming regions are rare in real rich and dense clusters of galaxies: the ambient very hot ($T \sim 10^8$ K) gas cannot cool, and thus cannot accrete onto dense lumps of collisionless stellar or dark matter. This effect sets in as soon as the strong caustics form in our simulations, and considerably before the final accumulation of virialized clusters.

We also sorted our computed sample into four sets by the metallicity of their first-generation stars ($Z/Z_\odot = < -1.69$, -1.69 to -1.57, -1.57 to -1.48, and > -1.48). The trends are comparable to those seen in Figures 2*a* and 2*b*, in the sense that galactic regions with initial high metallicity are more strongly biased than are those with initial low metallicity, and any set of objects has an increasing bias with decreasing scale. This presumably reflects the physical situation that high-density regions generally have higher metallicity. However, the trend is substantially weaker than that seen in subsets ordered by their formation time.

In Figure 3 we compare the simulation results with observations. In Figure 3*a* we select the second oldest quartile of objects and designate it as "elliptical galaxies," whose correlation function is shown along with the correlation functions of observed elliptical galaxies from two independent redshift surveys. Figure 3*b* shows the correlation function for the second youngest quartile of objects, which we call "spiral galaxies;" also shown are the correlation functions of observed spiral galaxies from redshift surveys. Finally, Figure 3*c* shows the correlation function of all galaxies and total mass compared with observed correlations of all galaxies from three redshift surveys. While the separate subsets (Figs. 3*a* and 3*b*) produce results in good agreement with their observational counterparts in the range of scales where our simulation results are most accurate (1–10 h^{-1} Mpc), the agreement between observations and simulations for all galaxies are not satisfactory. The likely cause is that, while we are computing the correlation functions of galaxy *mass* density, the observed correlation functions are galaxy number weighted, and in high-density regions the number density of objects will decrease due to mergers, while leaving the mass density in galaxies invariant. More robust is the trend found in our computations that older galaxies tend to cluster more strongly than average in a fashion that is fully consistent with observations.

Figure 4 shows the void probability function as a function of top-hat sphere radius. The void probability function is defined as the probability of a (top-hat) sphere of the given radius having a density of one-fifth of its corresponding global mean (Weinberg & Cole 1992). We see that the probability of having a void as large as 10 h^{-1} Mpc in the mass distribution (Fig. 4, *dot-dashed curve*) is 0.005 (the limited box size of the simulation does not allow a probability of less than 0.005 to be estimated). This has been highlighted by Peebles (1993) as a potential problem for CDM theories. However, the void probabilities for various subsets of galaxies and all galaxies are much larger, at $\sim$0.2–0.6, at $r_{\rm TH} = 10\ h^{-1}$ Mpc. Here we see that the simulations quite

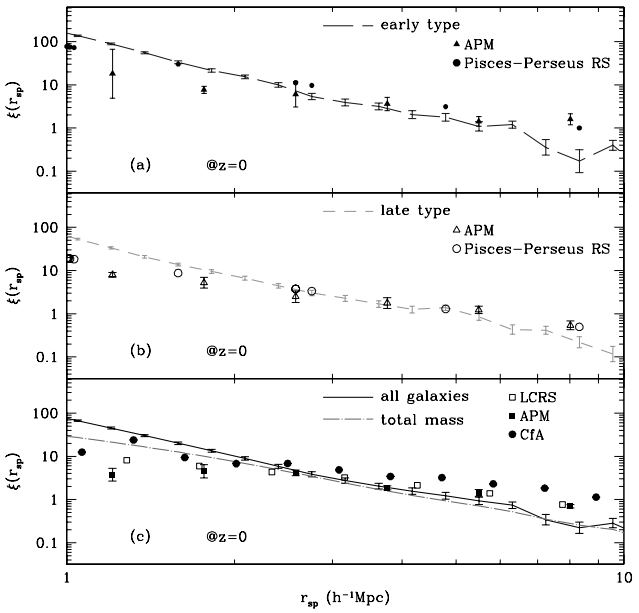

FIG. 3.—(*a*) Autocorrelation function for the early-type galaxies from the simulation (*long-dashed curve with 1 σ error bars*) at zero redshift; the second-oldest quartile of galaxies is used here. Also shown in (*a*) are observations for ellipticals from an analysis of galaxies in the Perseus-Pisces region (*filled circles*; Guzzo et al. 1997) and from the APM survey (*solid triangles with 1 σ error bars*; Loveday et al. 1995). (*b*) Autocorrelation function for the late-type galaxies from the simulation (*dashed curve with 1 σ error bars*) at zero redshift; the second-youngest quartile of galaxies is shown here. The symbols in (*b*) are observed correlations of spirals (*open circles*; Guzzo et al. 1997) in the Perseus-Pisces region and from the APM survey (*open triangles with 1 σ error bars*; Loveday et al. 1995). (*c*) Autocorrelation function for all galaxies (*solid curve*) and total mass (*dot-dashed curve*) at zero redshift. The symbols in (*c*) are observed correlation functions of galaxies from various surveys: *open squares*: from the Las Companas Redshift survey (Tucker et al. 1997); *filled squares*: from the APM survey (Loveday et al. 1995); *filled circles*: from the CfA survey (Vogeley et al. 1993).

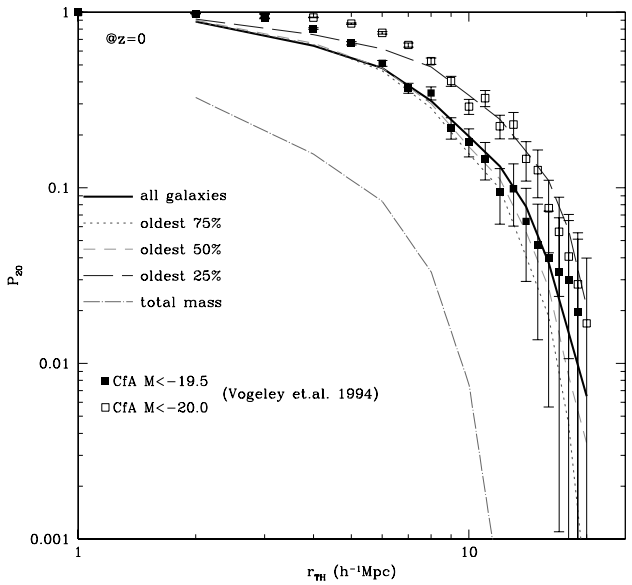

FIG. 4.—Void probability function as a function of top-hat sphere radius for all mass, all stellar mass, and three subsets of stellar mass (25%, 50%, and 75% oldest stellar mass). The void probability function is defined as the probability of a sphere of the given radius having a density of one-fifth of its corresponding global mean (Weinberg & Cole 1992). Also shown are observations from Vogeley et al. (1994) of the CfA survey for volume-limited samples with absolute magnitude less than −19.5 (*filled squares*) and −20.0 (*open squares*), respectively.

naturally produce a universe with over half of its volume in voids with radii greater than 5 h^{-1} Mpc, while only about 10% of the dark matter universe is in similar voids. Also shown as symbols are observations from Vogeley et al. (1994) of the CfA survey for volume-limited samples with absolute magnitude less than −19.5 (Fig. 4, *filled squares*) and −20.0 (*open squares*), respectively (fainter samples are not shown because our simulations are likely to have underproduced a significant number of fainter galaxies). Comparing with observations from Vogeley et al. (1994), we see that the real galaxies are indeed quite "voidy," consistent with the distribution of our simulated galaxies, but inconsistent with the distribution of mass, implying the necessity of biased galaxy formation. Also consistent with observations are the common trend that older galaxies are more voidy than younger ones, consistent with the fact that they tend to reside in high-density regions, such as rich clusters of galaxies. In the underdense regions examined in the analyses of voids, our spatial resolution is adequate to identify luminous galaxies.

Figure 5 shows the peculiar-velocity distribution of galaxy mass at $z = 0$, averaged over a top-hat smoothing scale of 1 h^{-1} Mpc, a typical group scale. Two things come to our attention. First, groups of younger galaxies tend to move more slowly than clusters of older galaxies, the difference of the median values between the youngest quartile and the oldest quartile being approximately 100 km s^{-1}. Second, averaging over all objects, stellar material tends to move slightly more slowly than dark matter. The second effect is equivalent to a small velocity antibias, reflecting the gasdynamic effect, while the first effect clearly indicates that older galaxies tend to reside in higher density, deeper potential regions than younger ones. The much larger effect, that the small-scale velocity dispersion of spirals is far less than the small-scale velocity dispersion of ellipticals, cannot be addressed in this work.

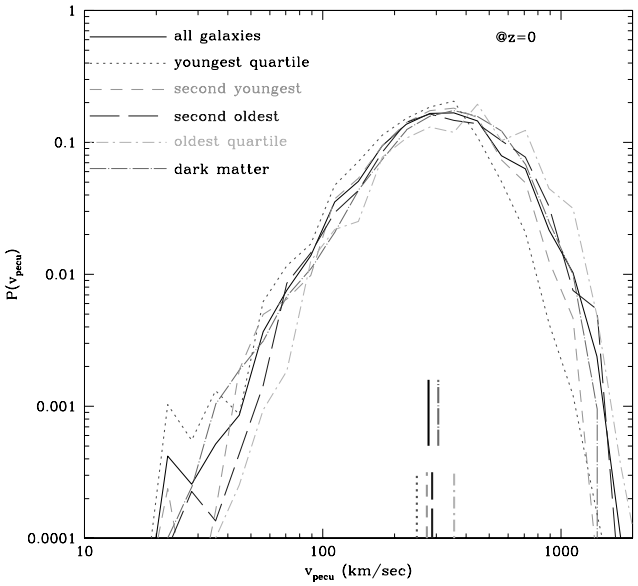

FIG. 5.—Peculiar-velocity distributions of various sets of galaxies and dark matter at $z = 0$ averaged over a top-hat smoothing scale of 1 h^{-1} Mpc. Vertical lines indicate the respective medians of the various sets of objects.

3.2. *Redshift Evolution of Galaxy Bias*

Let us now study bias at high redshift. Figure 2*c* shows the bias as a function of radius at four different redshifts. It can be seen that *bias is a monotonically increasing function of redshift*. Second, *bias is a monotonically decreasing function of scale at any redshift, with the rate of change as a function of smoothing scale* (db/dr_{cm}) *being nearly independent of redshift*. We think that the increase of bias with redshift is due primarily to the Kaiser (1984) effect. It has also been found in the simulations of Katz et al. (1999). Rarer and rarer events are required at higher and higher redshifts to produce a galaxy of a given mass. Finally, we predict that galaxies should be biased over mass by a factor of 3–4 on a $\sim 10\ h^{-1}$ Mpc scale at $z \sim 3$, in excellent agreement with recent observations of high-redshift galaxies at $z \sim 3$ (Steidel et al. 1998), which show strong concentrations of galaxies in narrow redshift bins and imply a strong bias of approximately the same magnitude. Note that our resolution at redshift $z = 3$ is $\sim 50\ h^{-1}$ kpc and $\sim 25\ h^{-1}$ kpc, respectively, in the two simulations

In addition, we have plotted the autocorrelation functions for all galaxy mass density and total mass density at various redshifts, as a function of *comoving* separation. We have found that, while the correlation of mass grows with time as predicted by linear theory, the correlation function of galaxy density as a function of *comoving* separation evolves very weakly (due to the countervailing evolution of bias) in the pair separation range $r \geq 2.5\ h^{-1}$ Mpc. This was also found by Katz et al. (1999) in a higher resolution simulation of a much smaller box ($11.1\ h^{-1}$ Mpc), which uses the same galaxy identification scheme as we do. Where our simulations overlap at redshift $z = 3$ and separation $r = 2\ h^{-1}$ Mpc, they obtain $\xi = 0.3$ and we find $\xi = 15.0$. The very significant difference is probably due to the differences in simulation box sizes.

In Figure 6 we address evolution in another way by examining the correlation length of the smoothed (mass-weighted) galaxy density as a function of redshift. Shown for comparison are the dark matter particle-particle (Fig. 6, *short-dashed curve*) and halo-halo (*long-dashed curve*) correlation lengths. Also shown as symbols and a shaded region are available observations from various sources. When needed, observations are converted to be consistent with the adopted cosmological model. We see that, as expected, the correlation function of mass (dark matter particles) is a monotonically deceasing function of redshift: gravitational instability (Peebles 1980) increases the correlation with time. On the other hand, the correlation of halos, selected by identifying all regions (cells) at or above the virial density at each redshift, is stronger at all redshifts than that of galaxy density. Furthermore, the halo correlation shows an interesting "dip" near $z \sim 0.5$, in contrast to the galaxy density correlation, which shows a weak peak at $z \sim 1.0$. The differences could be explained in part by numerical resolution effects and in part by real physical effects. Relatively larger dark matter particle masses compared to the baryonic mass resolution element in the simulation preferentially biases against identification of less massive halos, and thus overestimates the correlation of halos at all redshifts. The other effect is physical. At lower redshift ($z \leq 1.0$), the galaxy formation preferentially avoids rich clusters of galaxies, simply because these sites are (as noted earlier) too hot to permit proper cooling and gravitational

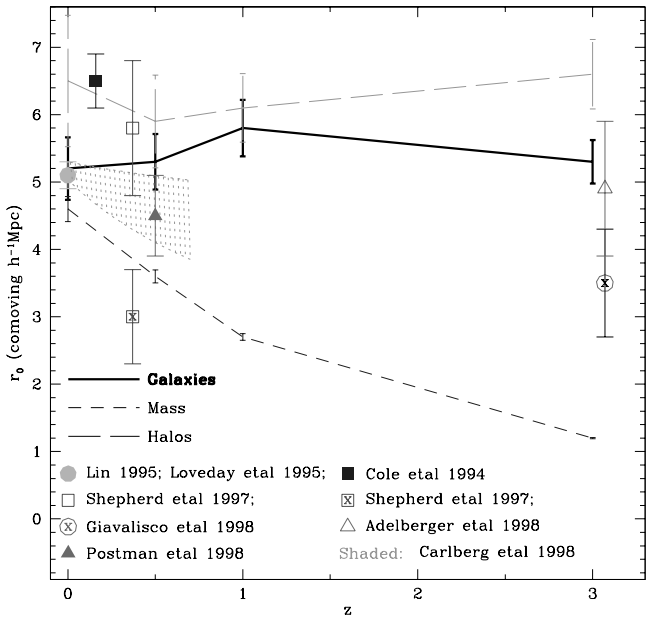

FIG. 6.—Correlation length of all galaxies (mass-weighted) as a function of redshift. Also shown are available observations from various sources: *filled circles*: APM observations (Loveday et al. 1995) and Las Companas Redshift survey (Lin 1995) at redshift zero; *filled square*: Autofib survey by Cole et al. (1994b) at $\bar{z} = 0.16$; *open squares*: Canadian Network for Observational Cosmology cluster survey by Shepherd et al. (1997), where the lower square with an "x" is from the extended field sample and the upper square includes the redshift subsample, at $\bar{z} = 0.37$; *hatched region*: result from CNOC2 by Carlberg et al. (1998) at $\bar{z} = 0.0$–0.7; *open circle with an* "x": Giavalisco et al. (1998) from analysis of Lyman-break galaxies at $\bar{z} = 3.07$; *open triangle*: Adelberger et al. (1998) from counts-in-cells analysis of Lyman-break galaxies at $\bar{z} = 3.07$; *filled triangle*: Postman et al. (1998) from an angular correlation analysis based on 710,000 galaxies with $I_{AB} < 24$ from a survey of 4×4 deg^2, contiguous sky. Also shown, for comparison, are the dark matter particle-particle (*short-dashed curve*) and halo-halo (*long-dashed curve*) correlation lengths.

instabilities to occur effectively. This explains the opposite trends in the halo and galaxy correlations at $z < 0.5$–1.0.

Kauffman et al. (1999) have also recently addressed this question and find a drop in the (comoving) correlation length at $z \sim 0.5$, very similar to what we found for dark matter halos, but contrary to what we found for our smoothed galaxy distribution. Hence, their result may be due to the method utilized, which is based on identifying N-body halos within which galaxies are expected to reside rather than attempting a full physical simulation. Thus, our result on the correlation evolution for halos (Fig. 6, *long-dashed curve*) is consistent with both that of Colin et al. (1999) from pure N-body simulations and that of Kauffmann et al. (1999) for gas-added halos in the general trend that the halo correlation decreases from redshift $z \geq 3$ to $z \sim 1$ and then rises towards zero redshift; there is a minimum at $z \sim 1$. We caution that the correlation at high redshift may have been overestimated in our simulation as a result of the limited resolution, which preferentially selects out the most massive halos that are most strongly clustered. This indeed may explain the somewhat stronger correlation found in our simulations than in Colin et al. (1999) at $z = 3$. At lower redshift, this effect is progressively less significant.

Overall, the fit to observations is reasonably good, both roughly consistent with being flat in the range $z = 0$–3. However, one needs to keep in mind that our computed correlation functions are weighted by the mass in galaxies,

not the number. Because of this, one expects to have better agreements between our results and observations that look at all galaxies and do not exclude high-density regions, if the cosmological model is right and the physical modeling is appropriate. Indeed, the top open square in Figure 6 of the Shepherd et al. (1997) data points, which does not exclude clusters in the sample, agrees much better with our results than the bottom open square with an "x," which does exclude clusters. The same can be said about the data point from Giavalisco et al. (1998; *open circle with an* "x"), which has a fainter sample than that of Adelberger et al. (1998; *open triangle*).

4. CONCLUSIONS

We analyze a new physically based large (100 h^{-1} Mpc) numerical simulation of a plausible ΛCDM cosmological model to quantify the relation expected between mass density and smoothed galaxy density. The results, while consistent, as far as we know, with existing observations, are quite inconsistent with the simplest model, where the ratio of galaxy fluctuations to mass fluctuations is a number, b, called bias. This ratio, in principle, could depend on the physical scale considered, the age or metallicity of the galaxies studied, or the epoch at which the analysis is made. We find in our simulation that bias increases with decreasing scale, with increasing galactic age or metallicity, and with increasing redshift. Looking at the average galaxy mass density at redshift zero, our result, $b = 1.35$, is consistent with the observed value 1.50 ± 0.19 (Loveday et al. 1996; the quoted observational value is for bright galaxies in the APM survey, which is appropriate for comparison with our model because the computed galaxy fluctuation is mass-weighted, not number-weighted, and is thus heavily weighted by the most massive, bright galaxies). All of the proposed dependencies can be tested by comparing to observations, and if confirmed can be used to interpret large-scale galactic surveys now underway. The surprising emptiness of voids also follows naturally from this physical treatment. We also predict that, due to an apparent coincidence between opposing trends, the strength of the spatial correlation between pairs of galaxies (as a function of *comoving* separation) should be essentially independent of redshift from $z = 0$ to $z = 3$. Here we differ from the Kauffmann et al. (1999) result, thus providing an opportunity to distinguish between the accuracy of the two methods currently being utilized for computing large-scale structure properties from ab initio methods. Finally, we find that the perturbations on the Hubble flow for young (e.g., late-type spiral or irregular) galaxies should be less than for early-type systems; this is a testable proposition.

Much of the complexity and stochasticity of the reported results is due, we believe, to the fact that the conventional treatments (including this one) are not examining all of the right variables. Blanton et al. (1999, 2000) show that galaxy density is a function not just of mass density, but also of potential (or velocity dispersion or temperature) in the region under examination. We will return to this in future work, with an analysis of the galaxy data using more appropriate variables.

Discussions with Michael Blanton, Guinevere Kauffmann, Jim Peebles, Michael Strauss, and Michael Vogeley are gratefully acknowledged, as are comments by a referee which enabled us to substantially clarify our presentation. We thank Michael Vogeley for providing the observations concerning void probability functions and correlation functions from the CfA Survey, Douglas Tucker for providing the correlation function data from the Las Campanas Redshift Survey, and Jon Loveday for providing the correlation function data from the APM Survey. This work is supported in part by grants NAG5-2759 and AST 93-18185, ASC 97-40300.

APPENDIX A

EFFECT OF FINITE NUMERICAL RESOLUTION

It is important to understand how the finite numerical resolution might have affected our computed results. In Figure 7 we show results from both $L = 100\,h^{-1}$ Mpc and $L = 50\,h^{-1}$ Mpc boxes. We smoothed the galaxy and dark matter density fields by a Gaussian window of the same radius, $0.5\,h^{-1}$ Mpc, in both boxes. Figure 7*a* shows the bias (defined by eq. [1]) from the

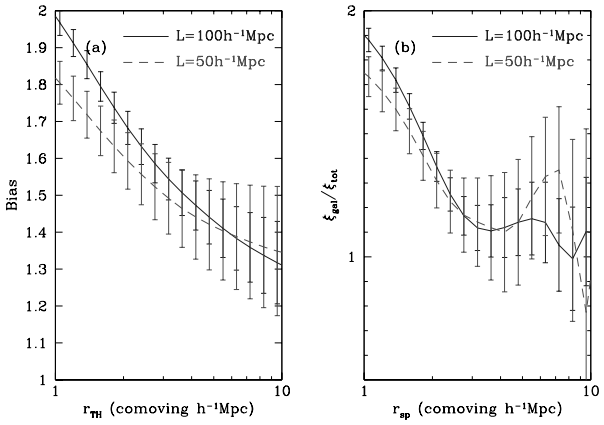

FIG. 7.—(*a*) Bias (defined by eq. [1]) from the two boxes with size $L = 100$ and $50\,h^{-1}$ Mpc. (*b*) Square root of the ratio of the galaxy correlation function to the total matter correlation function. The galaxy and dark matter density fields in both boxes are smoothed by a Gaussian window of radius $0.5\,h^{-1}$ Mpc. The error bars shown are 1 σ.

two boxes, and Figure 7*b* shows the square root of the ratio of the galaxy correlation function to the total matter correlation function. We see that the two boxes give comparable results, with a maximum difference of less than 9% (at $r_{TH} = 1\ h^{-1}$ Mpc, below which no comparison is made; the oscillations at $r_{sp} > 5\ h^{-1}$ Mpc in Fig. 7*b* are just noise). We therefore conclude that, for the scale range of interest, $r \geq 1\ h^{-1}$ Mpc, our finite numerical resolution does not significantly affect the results, at least for the quantities examined in this paper, namely, the bias of the smoothed galaxy density and the correlation function of this same function.

REFERENCES

Abel, T., Anninos, P., Norman, M. L., & Zhang, Y. 1998, ApJ, 508, 518
Adelberger, K., Steidel, C. C., Giavalisco, M., Dickinson, M. E., Pettini, M., & Kellogg, M. 1998, ApJ, 505, 18
Blanton, M., Cen, R., Ostriker, J. P., & Strauss, M. A. 1999, ApJ, 522, 590
Blanton, M., Cen, R., Ostriker, J. P., Strauss, M. A., & Tegmark, M. 2000, ApJ, 531, 1
Carlberg, R. G., et al. 1998, preprint astro-ph/9805131
Cen, R. 1992, ApJS, 78, 341
Cen, R., & Ostriker, J. P. 1992, ApJ, 399, L113
———. 1993, ApJ, 417, 415
———. 1999, ApJ, 517, 31
Cen, R., Phelps, S., Miralda-Escudé, J., & Ostriker, J. P. 1998, ApJ, 496, 577
Cole, S., Aragon-Salamanca, A., Frenk, C. S., Navarro, J., & Zepf, S. E. 1994a, MNRAS, 271, 781
Cole, S., Ellis, R., Broadhurst, T., & Colless, M. 1994b, MNRAS, 267, 541
Colin, P., Klypin, A. A., Kravtsov, A. V., & Khokhlov, A. M. 1999, ApJ, 523, 32
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
Eke, V. R., Cole, S., & Frenk, C. S. 1996, MNRAS, 282, 263
Frenk, C. S., White, S. D. M., Efstathiou, G., & Davis, M. 1985, Nature, 371, 595
Giavalisco, M., Steidel, C. C., Adelberger, K. L., Dickinson, M. E., Pettini, M., & Kellogg, M. 1998, ApJ, 503, 543
Gnedin, N. Y., & Ostriker, J. P. 1997, ApJ, 486, 581
Guzzo, L., Strauss, M. A., Fisher, K. B., Giovanelli, R., & Haynes, M. P. 1997, ApJ, 489, 37
Harten, A. 1983, J. Comput. Phys., 49, 357
Kaiser, N. 1984, ApJ, 284, L9
Kang, L., et al. 1994, ApJ, 430, 83
Katz, N., Hernquist, L., & Weinberg, D. H. 1992, ApJ, 399, L109
———. 1999, ApJ, 523, 463
Katz, N., Weinberg, D. H., & Hernquist, L. 1996, ApJS, 105, 19
Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, MNRAS, 307, 529
Kauffmann, G., Nusser, A., & Steinmetz, M. 1997, MNRAS, 286, 795
Kennicutt, R. C. 1989, ApJ, 344, 685
Krauss, L., & Turner, M. S. 1995, Gen. Relativ. Gravitation, 27, 1137
Kravtsov, A., & Klypin, A. 1999, ApJ, 520, 437
Lin, H. 1995, Ph.D. thesis, Harvard Univ.
Loveday, J., Efstathiou, G., Maddox, S. J., & Peterson, B. A. 1996, ApJ, 468, 1
Loveday, J., Maddox, S. J., Efstathiou, G., & Peterson, B. A. 1995, ApJ, 442, 457
Madau, P. 1997, in AIP Conf. Proc. 393, Star Formation Near and Far, ed. S. S. Holt & G. L. Mundy (New York: AIP), 481
Ostriker, J. P., & Steinhardt, P. 1995, Nature, 377, 600
Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton Univ. Press)
———. 1993, Principles of Physical Cosmology (Princeton: Princeton Univ. Press)
Pen, U.-L. 1998, ApJ, 504, 601
Postman, M., Lauer, T. R., Szapudi, I., &, Oegerle, W. 1998, ApJ, 506, 33
Ryu, D., Ostriker, J. P., Kang, H., & Cen, R. 1993, ApJ, 414, 1
Salaris, M., Degl'Innocenti, S., & Weiss, A. 1997, ApJ, 479, 665
Scherrer, R. J., & Weinberg, D. H. 1998, ApJ, 504, 607
Schmidt, M. 1959, ApJ, 129, 243
Schmidt, B. P., et al. 1998, ApJ, 507, 46
Serle, L., & Sargent, W. L. W. 1972, ApJ, 173, 25
Shepherd, C. W., Carlberg, R. G., Yee, H. K. C., & Ellingson, E. 1997, ApJ, 479, 82
Steidel, C. C., Adelberger, K. L., Dickinson, M., Pettini, M., Giavalisco, M., & Kellogg, M. 1998, ApJ, 492, 428
Steinmetz, M. 1996, MNRAS, 278, 1005
Trimble, V. 1997, in The Extragalactic Distance Scale, ed. M. Livio, M. Donahue, & N. Panagia (Cambridge: Cambridge Univ. Press), 407
Tucker, D. L., et al. 1997, MNRAS, 285, L5
Tytler, D., Fan, X.-M., & Burles, S. 1996, Nature, 381, 207
Weinberg, D. H., & Cole, S. 1992, MNRAS, 259, 652
White, S. D. M., Davis, M., Efstathiou, G., & Frenk, C. S. 1987, Nature, 330, 451
Vogeley, M. S. 1993, Ph.D. thesis, Harvard Univ.
Vogeley, M. S., Geller, M. J., Park, C., & Huchra, J. P. 1994, AJ, 108, 745