THE ASTROPHYSICAL JOURNAL, 517:54–63, 1999 May 20


# $N$-POINT CORRELATIONS IN CDM AND ΩCDM SIMULATIONS

ISTVÁN SZAPUDI,[1] THOMAS QUINN, JOACHIM STADEL,[2] AND GEORGE LAKE[2]

*Received 1998 October 12; accepted 1998 December 29*

## ABSTRACT

Higher order statistics are investigated in Ω cold dark matter (CDM) universes by analyzing 500 $h^{-1}$ Mpc high-resolution tree $N$-body simulations with both $\Omega = 1$ and $\Omega < 1$. The amplitudes of the $N$-point correlation functions are calculated from moments of counts-in-cells determined by a pair of new algorithms especially developed for large simulations. This approach enables massive oversampling with $\simeq 10^9$–$10^{14}$ cells for accurate determination of factorial moments from up to 47 million particles in the scale range of 8 $h^{-1}$ kpc–125 $h^{-1}$ Mpc. Thorough investigation shows that there are three scale ranges in the simulations: $\geq 8$ $h^{-1}$ Mpc, a weakly nonlinear regime where perturbation theory applies with utmost precision; 1–8 $h^{-1}$ Mpc, the nonlinear plateau; and finally $\leq 1$ $h^{-1}$ Mpc, a regime where dynamical discreteness effects dominate the higher order statistics. In the physically relevant range of 1–125 $h^{-1}$ Mpc the results (1) confirm the validity of perturbation theory in the weakly nonlinear regime; (2) establish the existence of a plateau in the highly nonlinear regime similar to the one observed in scale-free simulations; (3) show extended perturbation theory to be an excellent approximation for the nonlinear regime; (4) find the time-dependence of the $S_N$'s to be negligible in both regimes; (5) in comparison with similar measurements in the Edinburgh-Durham Southern Galaxy Catalog survey, strongly support $\Omega < 1$ with no biasing; and (6) show that the formulae of Szapudi and Colombi provide a good approximation for errors on higher order statistics measured in $N$-body simulations.

*Subject headings:* cosmology: theory — galaxies: clusters: general — large-scale structure of universe — methods: data analysis — methods: laboratory

## 1. INTRODUCTION

According to popular theories of structure formation, the distribution of mass in the universe grows by gravity from initially Gaussian fluctuations. The resulting distribution is described in a statistical way, most importantly via two-point and higher order correlation functions, which can be studied theoretically using analytical methods or numerical experiments. Although the comparison of the results with observations is somewhat complicated by the fact that galaxies do not necessarily trace mass (biasing), the manifold information contained in the higher order correlations in principle enables the separation of gravitational amplification from other processes (e.g., Fry 1994; Matarrese, Verde, & Heavens 1997; Scoccimarro et al. 1998; Frieman & Gaztañaga 1999; Scoccimarro, Szapudi, & Frieman 1999; Szapudi 1998).

Following the pioneering work of Peebles and collaborators (e.g., Fry & Peebles 1978; Peebles 1980 and references therein), perturbation theory (PT) became the prime analytical tool to study higher order correlation functions. The Euler equations for a gravitating fluid are expanded around small fluctuations to predict the amplitudes of the correlation functions at weakly nonlinear scales. In contrast, $N$-body simulations calculate the gravitational amplification directly; thus, up to numerical accuracy, they follow the full nonlinear evolution. Simulations not only yield beautiful agreement with PT at large scales, but they also penetrate the highly nonlinear evolution of smaller scales. These scales are especially important, since, except for the largest galaxy catalogs, most observations are performed at small or intermediate scales. The method of moments of counts in cells is especially useful for comparison, since the moments were calculated in the framework of PT (e.g., Peebles 1980; Juszkiewicz, Bouchet, & Colombi 1993; Bernardeau 1992; Bernardeau 1994; Bernardeau 1995) and measured in $N$-body simulations (e.g., Bouchet, Schaeffer, & Davis 1991; Bouchet, & Hernquist 1992; Baugh, Gaztañaga, & Efstathiou; Gaztañaga, & Baugh 1995; Colombi, Bouchet, & Hernquist 1995) and galaxy catalogs as well (e.g., Peebles 1980; Gaztañaga 1992; Szapudi, Szalay, & Boschán 1992; Meiksin, Szapudi, & Szalay 1992; Bouchet et al. 1993; Gaztañaga 1994; Szapudi et al. 1995; Szapudi, Meiksin, & Nichol 1996; Kim, & Strauss 1998; Szapudi & Szalay 1997a).

While the simplest version of standard cold dark matter initial conditions appears to be excluded by observations of the variance as measured by the cosmic microwave background, cluster abundances, pair-wise velocities, and galaxy clustering, it is qualitatively the most successful theory, against which every other theory is measured. In this work large, high-resolution cold dark matter (CDM) simulations are used in an attempt to understand clustering with unprecedented errors in a large dynamic range. Motivated by observations, a low-density variant of CDM (ΩCDM) is investigated as well, since it is one of the most viable alternatives at present.

Moments of counts in cells are used to quantify higher order clustering in the simulations. Similar previous measurements are improved upon in several ways: a large 500 $h^{-1}$ Mpc box size is used to diminish finite volume effects, i.e., the error on the measurement from fluctuations of the universe on scales larger than the box size; $47 \times 10^6$ particles are used for a large dynamic range; a pair of new methods are employed calculate counts in cells, which are especially designed for large simulations and to minimize the measurement errors; for quantitative assessment of the

[1] University of Durham, Department of Physics, South Road, Durham DH1 3LE, England, UK.
[2] Department of Astronomy, University of Washington, Seattle, WA 98195-1580.

accuracy a strict theoretical error analysis is performed according the formalism of Szapudi & Colombi (1996, hereafter SC) and Szapudi, Colombi, & Bernardeau (1999, hereafter SCB). Because of the above properties, the measurements are relevant to study both the highly and mildly nonlinear regimes as well as the transition between them. Special care is taken to determine the scales of reliability, and appropriate tests are done to estimate the artificial two-body relaxation effects, which appear to be the limiting factor at small scales.

The organization of the paper is as follows: The next section outlines the method of counts in cells as used here, § 3 describes the simulations and establishes the scales of reliability, § 4 presents the measurements of the cumulants in the various simulations, and § 5 discusses findings in terms of PT and extended perturbation theory (EPT) providing an efficient framework to compress the results and facilitating the comparison with observational data from the Edinburgh-Durham Southern Galaxy Catalog (EDSGC) survey. The Appendix contains the definition of the pair of algorithms used to calculate counts in cells.

## 2. METHOD

A substantially improved version of the counts in cells method is used in this work.[3] It consists of calculating the amplitudes of higher order correlation functions in a sequence of three consecutive steps: estimation of the probability distribution; calculation of the factorial moments; and extraction of the normalized, averaged amplitudes of the $N$-point correlation functions, the $S_N$'s. The relevant definitions and theory are briefly summarized below, while Szapudi et al. 1996 and references therein can be consulted for more details.

Let $P_N$ be the probability that a randomly thrown cell in the simulation contains $N$ particles, with implicit dependence on the cell size $\ell$. The estimator for this is the frequency distribution

$$\tilde{P}_N = \frac{1}{C}\sum_{i=1}^{C}\delta(N_i = N)\,, \tag{1}$$

where $C$ is the number of cells thrown and $N_i$ is the number of objects in cell $i$. It is desirable to use as many cells as possible, since for large $C$ the measurement errors associated with the finite number of cells behave as $C^{-1}$ (SC). Here the main improvement over the more traditional approach is the pair of algorithms described in the Appendix, which enable us to use $C \simeq 10^9$–$10^{14}$ even in these large simulations.

The factorial moments (see, e.g., Szapudi & Szalay 1993) may be obtained from the probability distribution using

$$F_k = \sum P_N (N)_k\,, \tag{2}$$

where $(N)_k = N(N-1)\ \ldots\ (N-k+1)$ is the $k$th falling factorial of $N$. The $F_k$'s directly estimate the moments of the hypothetical continuum random field, which is Poisson sampled by the simulation particles. This is the most accurate and efficient way of subtracting shot noise, which becomes important on small scales. Note that for estimation purposes the estimator of the probability distribution is substituted in the above equation; i.e., $\tilde{P}_N \to P_N$.

The average of the $N$-point angular correlation functions on a scale $\ell$ is defined by

$$\bar{\xi}_N(\ell) = V^{-N}\int dV_1\ \ldots\ dV_N\,\xi_N(r_1, \ldots, r_N)\,, \tag{3}$$

where $\xi_N$ is the $N$-point correlation function in the simulations and $V$ is the volume of a cell. We define $S_N$ in the usual way,

$$S_N = \frac{\bar{\xi}_N}{\bar{\xi}_2^{N-1}}\,. \tag{4}$$

The factorial moments have an especially simple relation to the $S_N$'s through the recursion relation (Szapudi & Szalay 1993), which is quoted for completeness:

$$S_k = \frac{F_k\bar{\xi}_2}{N_c^k} - \frac{1}{k}\sum_{q=1}^{k-1}\frac{1}{N_c^q}\left[(k-q)S_{k-q}F_q\binom{k}{q}\right], \tag{5}$$

where $N_c = \langle N\rangle\bar{\xi}_2$.

The most critical and CPU-intensive component of the above procedure is the calculation of counts in cells with appropriate oversampling. While there exists an algorithm for infinitely oversampling by Szapudi 1997, it would be impractical for 47 million particles in three dimensions. Therefore a new approach was developed especially for large simulations; the resulting pair of algorithms for smaller and larger scales have substantial overlap at intermediate scales suitable for testing. They are detailed in the Appendix. With a modest 6–8 hr of CPU investment, these algorithms can achieve $C \simeq 10^9$–$10^{14}$ sampling cells simultaneously at a hierarchy of scales between 1/65536 and 1/4 of the simulation box size.

## 3. MEASUREMENTS

### 3.1. *Simulations*

The characteristics of the simulations used are summarized in Table 1. The box size, particle number, and force-softening of the large simulations were chosen to model the formation of galaxy clusters in a volume of the universe comparable to that to be surveyed by the Sloan Digital Sky Survey (SDSS; Gunn & Knapp 1993) All simulations were computed using PKDGRAV (Stadel & Quinn 1999), a scalable parallel treecode with periodic boundary conditions. Accurate forces were maintained by using a cell opening angle of $\theta = 0.8$ for $z < 2$ and $\theta = 0.6$ for $z > 2$ and by expanding the potentials of cells to hexadecapole order. Time steps were constrained to $\Delta t < 0.3(\epsilon/v_{\max})$, where $\epsilon$ is the softening length and $v_{\max}$ is the approximate maximum speed. A cubic spline softening kernel was used. The simulations were started at $z = 49$ for the $\sigma_8 = 1$ models; thus the transients from initial conditions should be negligible (Scoccimarro 1998). The same simulations were also used by Governato et al. 1999 to explore the properties of galaxy clusters. Note that $I_a$, $I_b$, and $I_c$ are the same simulation at different output times; likewise, $II_a$, $II_b$, and $II_c$ are as well. Simulations $i_{c1}\ \ldots\ i_{c5}$ are an ensemble of simulations with different realizations of the same initial power spectrum.

### 3.2. *Scales of Reliability*

Since both algorithms in the Appendix employ powers of 2, initially the scale range of $2^{-16}\ \ldots\ 2^{-2}$ times box size was used for calculating counts in cells, corresponding to 7.63

[3] Note that in an $N$-body simulation edge effects are eliminated by the periodic boundary conditions; therefore the edge-corrected estimator of Szapudi & Szalay 1998 is not necessary.

TABLE 1

ANALYSIS OF SIMULATIONS

| Label | $\Omega$ | $H_0$ (km s$^{-1}$ Mpc$^{-1}$) | $\sigma_8$ | $L_{box}$ ($h^{-1}$ Mpc) | $N$ | $\epsilon$ ($h^{-1}$ kpc) | $l_{50}$ ($h^{-1}$ Mpc) |
|---|---|---|---|---|---|---|---|
| $I_a$ ..... | 1.0 | 50 | 0.5 | 500.0 | $4.7 \times 10^7$ | 50.0 | 4 |
| $I_b$ ..... | 1.0 | 50 | 0.7 | 500.0 | $4.7 \times 10^7$ | 50.0 | 2 |
| $I_c$ ..... | 1.0 | 50 | 1.0 | 500.0 | $4.7 \times 10^7$ | 50.0 | 1 |
| $II_a$ ..... | 0.5 | 75 | 0.74 | 500.0 | $4.7 \times 10^7$ | 50.0 | 3 |
| $II_b$ ..... | 0.4 | 75 | 0.88 | 500.0 | $4.7 \times 10^7$ | 50.0 | 1.3 |
| $II_c$ ..... | 0.3 | 75 | 1.0 | 500.0 | $4.7 \times 10^7$ | 50.0 | $\leq 1$ |
| $i_{c0}$ ..... | 1.0 | 50 | 1.0 | 500.0 | $3.0 \times 10^6$ | 160.0 | 13 |
| $i_{c1}$ ... $i_{c5}$ ..... | 1.0 | 50 | 1.0 | 200.0 | $3.0 \times 10^6$ | 50.0 | 1 |

NOTE.—Values are as follows: $\Omega$ is the density parameter, $H_0$ is the Hubble constant, $\sigma_8$ is the RMS density fluctuation in 8 $h^{-1}$ Mpc spheres, $L_{box}$ is the size of the periodic box, $N$ is the number of particles, and $\epsilon$ is the force softening length. Note that $i_{c1}$ ... $i_{c5}$ are five different realizations of the same initial conditions.

$h^{-1}$ kpc–125 $h^{-1}$ Mpc. The lower scale is smaller than the softening length used for force calculation. With our algorithm (A2) for small scales, we could obtain almost arbitrarily small scales for free. The upper scale still contains 256 nonoverlapping volumes, sufficient for a distribution not far from Gaussian. Figure 1 shows the counts in cells distribution for simulation $I_c$. The curves from right to left correspond to scales of 1/4, 1/8, ... , 1/65536$L$ (the box size). Algorithm (A1) was used for scales down to 1/512$L$, and algorithm (A2) for smaller scales. Algorithm (A1) uses a fixed number of cells $C \simeq 1.1 \times 10^9$ for all scales, while algorithm (A2) increases $C$ eightfold at each step toward smaller scales after starting with the above value. This is reflected in Figure 1 by the lowest possible value $P_N$ can take. The tail of the distribution still shows some wavering, which could be smoothed out with even higher oversampling. The next section, however, will show that the resulting measurement errors are much smaller than the theoretical variance of the simulations; therefore the sampling is sufficient. The high degree of oversampling was made possible only by the algorithms of the Appendix specifically developed for this purpose.

The upper panel of Figure 2 shows the variance, or the average two-point correlation function over a cell, as calculated from the first two factorial moments of simulation $I_c$. The solid curve is algorithm (A1), and the joining dashed line shows algorithm (A2). The triangles and squares display the expected variance in cubic windows obtained by integrating the linear and nonlinear power spectrum, respectively [Peacock & Dodds 1996; the nonlinear $P(k)$ was provided by C. Baugh 1998, private communication]. The fitting formula for the nonlinear power spectrum provides a good approximation, the largest discrepancy being roughly 30% on the smallest scales. The lower resolution simulation, discussed later (Fig. 2, *dotted line*), is in even better agreement with the fitting formula. While providing a more accurate fitting formula for high-resolution simulations could be a topic of further investigation, this work concen-

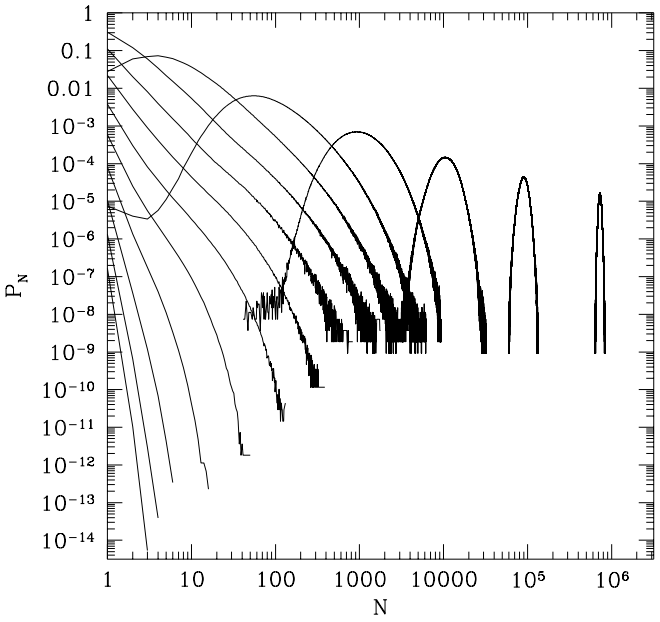


FIG. 1.—Counts in cells distribution for simulation $I_c$ (see text) from scales of 7.63 $h^{-1}$ kpc doubling up to 125 $h^{-1}$ Mpc from left to right. The scale corresponds to the size of the cubical window. Algorithm (A1) was used for scales $\geq 1$ $h^{-1}$ Mpc, and algorithm (A2) was used for the smaller scales. Note that the accuracy is as much as $10^{-14}$ and at least $10^{-9}$.

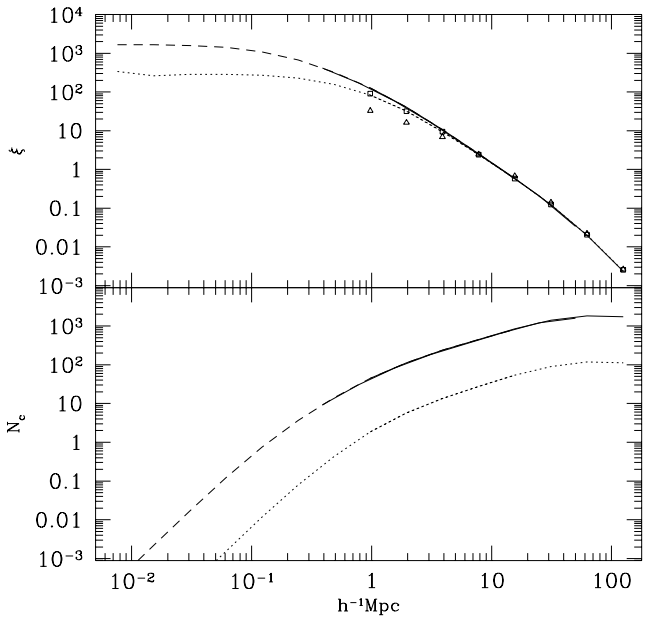


FIG. 2.—*Upper panel*: Joining solid and dashed lines displaying the measured variance $\bar{\xi}$ in cubical cells as a function of scale as calculated with algorithms (A1) and (A2) (*solid and dashed lines*, respectively). The triangles and squares are the predictions from the linear and nonlinear power spectra, respectively. The dotted line is the same for simulation $i_0$. *Lower panel*: $N_c = \bar{\xi}\bar{N}$ as a function of scale. The line types are the same as for the upper panel.

trates on higher order statistics; the second-order moment is only shown as a test.

The higher order factorial moments were calculated as well from the counts in cells according § 2. The resulting $S_N$'s up to ninth order are displayed on Figure 3. Again, the solid and dashed lines are the results from the two algorithms of the Appendix for small and large scales. Note the excellent agreement in the overlap, despite the fact that the sampling is somewhat different because of the random shifts employed in algorithm (A2).

Qualitatively, one can distinguish three regimes in Figure 3. The dot-dashed lines display theoretical predictions from PT (Juszkiewicz et al. 1993; Bernardeau 1994) up to sixth order. The agreement is excellent from scales upward of 8 $h^{-1}$ Mpc, in the weakly nonlinear regime. Between 1–8 $h^{-1}$ Mpc in the highly nonlinear regime, the $S_N$'s are higher than the PT prediction because of enhanced nonlinear effects. They constitute a shallow plateau in agreement with previous results from scale invariant simulations (Colombi et al. 1995). Finally, downward from 1 $h^{-1}$ Mpc there seems to be a third regime with a steeper rise. As illustrated next, this is caused by artificial particle discreteness effects.

The lower panel of Figure 2 plots $N_c = \bar{\xi}\bar{N}$, the number of particles in a typical cluster, as a function of scale; $N_c$ indicates how well the simulation represents the fluid limit. For small values the dynamics in typical clusters is artificially dominated by particle discreteness effects, a dynamical shot noise (Colombi et al. 1995). Such effects do not represent real physics, since particles in the simulation should follow the dynamics of the underlying smooth field. Indeed, at scales smaller than 1 $h^{-1}$ Mpc, $N_c$ becomes fairly small, which likely explains the sharp rise in the $S_N$'s.

To test this idea, several auxiliary CDM simulations were run with 3 million particles, and the $S_N$'s were measured. Simulation $i_{c0}$ had the same initial conditions and box size as the main simulation; therefore the shot noise is expected to turn up the $S_N$'s on larger scales, if the above explanation is correct. According to the dotted line on the lower panel of Figure 2, which displays $N_c$ for $i_{c0}$, the break is expected to happen at around 10–15 $h^{-1}$ Mpc, if the suspected scaling with $N_c$ is correct. Indeed, this seems to be the case for the dotted lines on Figure 3, supporting the role of particle discreteness in the artificial increase of the higher order cumulants.

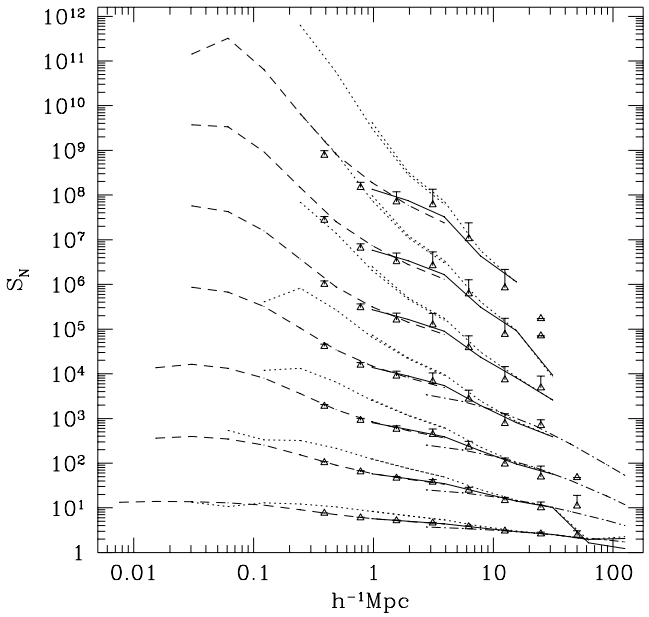


FIG. 3.—Cumulants $S_N$ displayed in increasing order upward up to ninth order as measured in simulation $I_c$ with algorithms (A1) and (A2) (*solid and dashed lines*, respectively). The dot-dashed lines show the predictions of PT up to sixth order. The dotted lines are similar measurements in simulation $i_0$. The triangles with error bars display the average and dispersion of measurements in simulations $i_1 \ldots i_5$. Only the upper error bar is plotted to reduce clutter.

For another test a set of five simulations were run, $i_{c1} \ldots i_{c5}$. They had smaller box size, $200^3$ $h^{-1}$ Mpc$^3$, to keep the average number of particles the same as the original simulation. The initial conditions were independently generated for each realization. The ensemble average of these simulations is expected to yield the same results upward from 1 $h^{-1}$ Mpc as the original simulation, perhaps with less accuracy because of the enhanced cosmic error caused by the smaller volume (SC). These measurements are displayed as triangles on Figure 3. The error bars were calculated by estimating the dispersion of the five simulations $i_{c1}$–$i_{c5}$; only the upper error bar is displayed for clarity. The results are in excellent agreement with the expectations, further supporting the idea that the third regime at small scales is a sign of dynamical discreteness effects.

Colombi et al. 1995 found that if $l_c$ is defined by $N_c(l_c) = 1$, a sufficient condition for the fluid limit is $l \geq 1.5 l_c$. The location of the break in the curves suggest a slightly more conservative limit in such a way that $N_c(l_{50}) = 50$. This somewhat ambiguous prescription depends on the details of the simulation and the desired precision of the agreement between the fluid limit and the measurements at each order. Our choice corresponds to 1 $h^{-1}$ Mpc as the scale of reliability. Note that the accuracy depends on the order, deteriorating toward the higher moments. It seems more logical to relax the required precision toward higher order than to define a set of scales of reliability becoming larger with higher order. This somewhat arbitrary but natural choice of 1 $h^{-1}$ Mpc is adopted for the measurements performed in the rest of the simulations, but $l_{50}$ is given for reference in Table 1 for each output. Note also that for the two-point correlation function only, a smaller $N_c$, and a correspondingly smaller scale, is sufficient (see, e.g., Jain 1997).

## 4. RESULTS

According to the previous reasoning, it is meaningful to extract higher order correlations down to $\simeq 1$ $h^{-1}$ Mpc only; thus only algorithm (A1) was sufficient for the rest of the measurements. Although the force resolution would suggest a lower threshold, as detailed above, particle discreteness effects raise the scale of reliability. Six outputs of two high-resolution CDM simulations summarized in Table 1 ($I_a$, $I_b$, $I_c$, $II_a$, $II_b$, $II_c$,) were used to measure counts in cells on scales in the range 1–125 $h^{-1}$ Mpc.

The measurements of the $S_N$'s, the main result of this paper, are displayed in Figures 4 with solid lines. The $\Omega = 1$ and $\Omega < 1$ simulations are displayed on the left-hand and right-hand sides, respectively. For reference, the measured $S_3$ and $S_4$ are given in Tables 2 and 3 as well. The weakly nonlinear regime on large scales is distinguished from the nonlinear plateau at small scales in all cases. The behavior of the higher order moments is qualitatively similar to scale-invariant simulations (Colombi et al. 1995). PT predicts that the $S_N$'s are independent of the output times. This appears to be a good approximation even in the highly

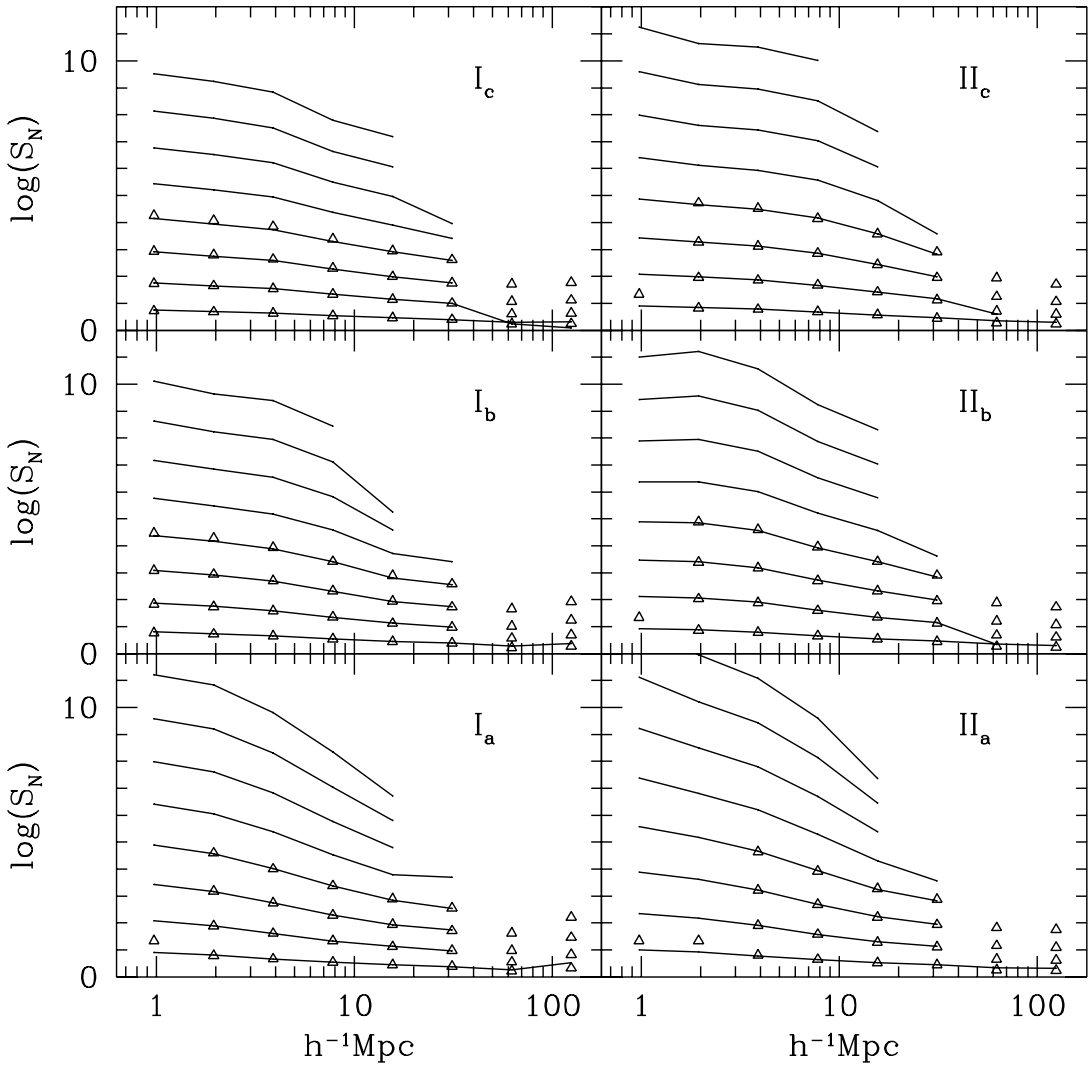


FIG. 4.—Measurements of $S_N$, $3 \leq N \leq 10$, in all simulations are plotted with solid lines, in increasing order upward. The triangles represent the $S_N$'s from EPT, when the best fit $n_{\rm eff}$ is used.

nonlinear regime, especially down to scales of $l_{50}$. For instance, for the $\Omega = 1$ simulations on 4 $h^{-1}$ Mpc ($l_{50}$ for $I_a$), $S_3$ changes only about 5%, which is the same order as the errors. Even $S_{10}$ is constant within a factor of 2–3; i.e., the higher order moments are constant within the errors (see next subsection). The decreasing trend on small scales can be explained by contamination effects from particle discreteness. As the scale of reliability moves to the left for the more relaxed, later simulations, the $S_N$'s decrease slightly. These initial observations will be refined by comparing with the predictions of PT and EPT in the next section, after the error budget is detailed in the next subsection.

TABLE 2

TABULATION OF $S_3$ MEASUREMENTS FOR DIFFERENT SIMULATIONS

| $l$ ($h^{-1}$ Mpc) | $I_a$ | $I_b$ | $I_c$ | $II_a$ | $II_b$ | $II_c$ |
|---|---|---|---|---|---|---|
| 0.976 ..... | 7.98 | 6.45 | 5.73 | 10.22 | 8.43 | 7.94 |
| 1.953 ..... | 6.41 | 5.57 | 5.00 | 8.49 | 7.60 | 7.05 |
| 3.9 ..... | 4.61 | 4.55 | 4.38 | 6.09 | 6.24 | 6.08 |
| 7.81 ..... | 3.47 | 3.54 | 3.56 | 4.30 | 4.54 | 4.78 |
| 15.625 ..... | 2.87 | 2.89 | 2.93 | 3.38 | 3.51 | 3.68 |
| 31.25 ..... | 2.43 | 2.47 | 2.52 | 2.87 | 2.94 | 2.99 |
| 62.5 ..... | 1.87 | 1.92 | 1.96 | 2.20 | 2.25 | 2.29 |
| 125 ..... | 3.34 | 2.40 | 2.07 | 2.07 | 2.03 | 2.00 |

NOTE.—The properties of simulations can be found in Table 1; $l$ is the size of the cubical window in which counts in cells where measured.

### 4.1. *Errors*

According to SC, the errors on the previous results can be classified into measurement errors and cosmic errors.

TABLE 3

TABULATION OF $S_4$ MEASUREMENTS FOR DIFFERENT SIMULATIONS

| $l$ ($h^{-1}$ Mpc) | $I_a$ | $I_b$ | $I_c$ | $II_a$ | $II_b$ | $II_c$ |
|---|---|---|---|---|---|---|
| 0.976 ..... | 121.1 | 74.8 | 58.1 | 222.7 | 133.6 | 121.5 |
| 1.953 ..... | 80.0 | 56.7 | 44.7 | 154.4 | 115.2 | 96.5 |
| 3.9 ..... | 41.4 | 38.7 | 34.7 | 80.7 | 80.8 | 74.1 |
| 7.81 ..... | 21.5 | 22.5 | 21.8 | 36.9 | 40.4 | 46.7 |
| 15.625 ..... | 13.4 | 13.6 | 14.2 | 20.1 | 22.5 | 25.7 |
| 31.25 ..... | 9.1 | 9.6 | 10.2 | 13.8 | 14.4 | 14.8 |

NOTE.—Same as Table 2 for $S_4$.

The measurement errors arise from a finite number of cells, $C$, being used to estimate the distribution of counts in cells. The appropriate expression for the error generating function is (see SC for details):

$$E(x, y) = (1/C)[P(xy) - P(x)P(y)] . \quad (6)$$

The expansion of this equation yields the measurement error in the $N$th moment, which depends on the $2N$th moment. If $C \to \infty$, the contribution approaches 0 as expected. We used this equation self-consistently to obtain errors up to fifth order (since tenth-order moments were measured). The measurement error is largest at the smallest scales. Figure 5 shows the relative measurement error as a function of order for the 1 $h^{-1}$ Mpc scale for $I_c$. Since there is a convex curvature on the graph, the continuing dotted line is a conservative overestimation of the errors for the orders $N > 5$. This suggests that even at tenth order the measurement errors contribute less than 10%; thus a further increase in the sampling is not required. This finding is true for the other simulations as well.

Cosmic errors are an inherent property of the simulations and cannot be improved upon except by using a larger volume or an ensemble of realizations. This type of error can be classified as finite volume, discreteness, and edge effects (SC). They arise, respectively, from the (hypothetical) fluctuations on scales larger than the simulation, the finite number of particles used to model the density field, and the uneven weighting of points. Because of the periodic boundary conditions, edge effects are not significant; neither are discreteness effects, except for the smallest scales because of the large number or particles used. Therefore finite volume effects are expected to be the dominant contribution if systematic errors from the inaccuracy of the calculations are not considered.

Two methods were used for estimating the cosmic errors: measuring the dispersion numerically from the ensemble of simulations $i_{c1}$ ... $i_{c5}$ (see the error bars in Fig. 3) and using the theory of (SC; SCB) to estimate the errors from the measured higher order moments self-consistently up to fourth order. The details of the calculations can be found there; here we only summarize the basic idea.

SC calculated the generating function of the variance of factorial moments due to edge, discreteness, and finite volume effects. Since the connected moments can be expressed in terms of the factorial moments (which are the discrete version of the disconnected moments), their results can be used to express the errors on the connected moments (see SCB for more details). The resulting formulae express the errors on the $N$th-order connected moments in terms of the $2N$th connected moments for $N \leq 4$. The expressions are too complicated to quote here (they are over 500 lines long); therefore only the self-consistent numerical estimates are used. For the case of the connected moments it is not possible to simply separate the different contributions for the errors. Therefore discreteness and edge effects are included in the calculations, even though this way the errors could be overestimated at large scales according to the previous considerations.

Figure 6 compares the numerical estimates of the theoretical error calculation (see also Colombi, Szapudi, & the Virgo Collaboration 1999) The solid line displays the unbiased estimate of the variance for $S_3$, and $S_4$, and the dotted lines show the theoretical calculation of the errors in the individual simulations $i_{c1}$ ... $i_{c5}$. The dashes are the result of a theoretical calculation as well but using the ensemble average of the five simulations for the moments. While the theoretical estimates from the individual simulations are in excellent agreement with the empirical dispersion, the average of the five simulations curiously overestimates the errors, especially on larger scales. A possible explanation is that since the error distribution is skewed (SC), a few overshoots can dominate the average.

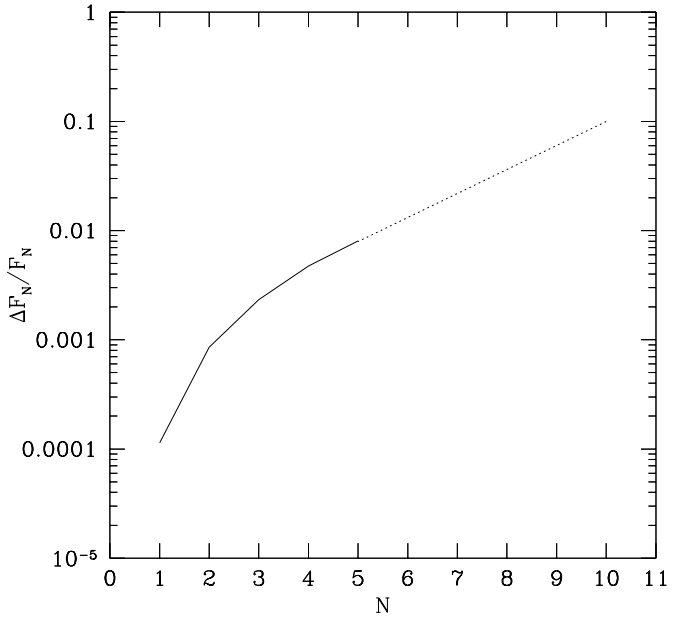


FIG. 5.—Relative measurement errors are plotted as a function of order. For $N \leq 5$ the errors were calculated with the formulae of SC. The dotted line is a conservative extrapolation of the calculations, which is likely to overestimate the actual errors.

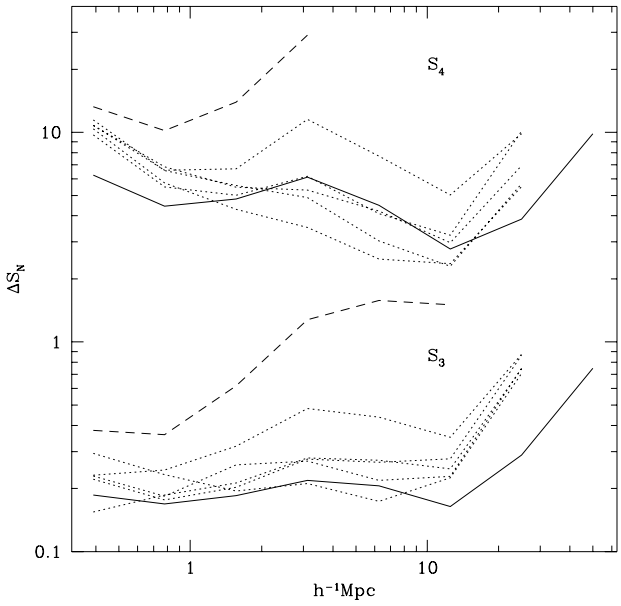


FIG. 6.—Absolute errors for $S_3$ and $S_4$ are plotted for simulations $i_1$ ... $i_5$. The lower set of curves represent $S_3$; the upper set represent $S_4$. The solid lines show the errors estimated by calculating the variance in the fives simulations. The dotted lines show the theoretical error for each of the five simulations calculated from the measured cumulants self-consistently. The dashed lines are the errors estimated from moments determined by the ensemble average of the five simulations.

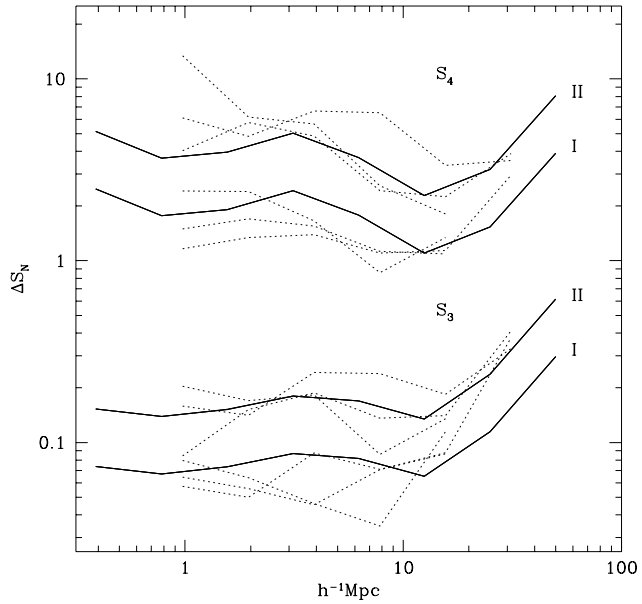


FIG. 7.—Theoretical errors for $S_3$ and $S_4$ are shown in simulations $I_a$, $I_b$, $I_c$, $II_a$, $II_b$, $II_c$. The lower set of curves refer to $S_3$; the upper set refer to $S_4$. Within each set there are two solid lines with approximately matching dotted lines. The upper quadruplet refers to simulations $II_*$; the lower one refers to $I_*$. The solid lines show the dispersion measured in simulations $i_1 \ldots i_5$ scaled with $2(\bar{\xi})^{1/2}$ of $I_c$ (*lower set*), and $II_c$ (*upper set*), respectively. The three dotted lines in each quadruplet display the different output times *a*, *b*, *c*.

This is amplified by the nonlinear expressions used to estimate the errors. The agreement nevertheless is surprisingly good, despite the anticipated "error on the error" problem (SC): the error on the fourth-order moment depends on eighth-order quantities, and the error on the error depends on up to sixteenth-order moments. To determine empirically the errors with negligible variance, 16 orders should be controlled with high precision, which is hardly possible using only five simulations of this size. For instance, in SC 1000 subsamples were needed to control the error on the error. Nevertheless, we can draw from the figure the conclusion that the theoretical calculations for the *individual* simulations are in excellent agreement with the empirical dispersion. We generalize this finding to the other simulations, where an ensemble of realizations is not presently available; i.e., we assume that the theoretical calculation is a good estimate of the errors up to fourth order. The error calculation yields less than 1% error for $S_3$ and about 5% for $S_4$ at most scales, except perhaps at the largest scales, where the errors appear to turn up to few tens of percent.

In fact, for simulation $I_c$, which has the exact same properties as simulations $i_{c1} \ldots i_{c5}$, except that it is larger, it is interesting to try the following naive scaling: if, as argued above, finite volume effects dominate, the errors on the disconnected moments scale with the variance over the full box, $(\xi L)^{1/2}$. Even though for the connected moments the formulae are more complicated, we find empirically that scaling with $2(\xi L)^{1/2}$ is an excellent approximation for the errors (see Fig. 7). Moreover, it appears that the earlier outputs have the same absolute error, i.e., the same scaling. Encouraged by these findings, a similar scaling was applied to the $\Omega < 1$ simulations. Again, scaling the $2(\xi L)^{1/2}$ of the last output is an excellent approximation. These observations are valid at the factor of 2 level: a considerable accuracy if the arguments about the "error on the error" are taken into account. We conjecture that similar approximations can be used at higher order.

## 5. DISCUSSION

The approximations developed for the errors in the previous section facilitate the comparison of the results with observations. The framework for comparison is naturally provided by PT and its generalization for smaller scales, EPT. PT gives simple expressions for the higher order correlation amplitudes $S_N$ at any order $N$. For instance for the third-order quantity $S_3 = 34/7 - (n + 3)$ (see, e.g., Juszkiewicz et al. 1993), where $n$ is the local index of the power spectrum. This formula, and the corresponding ones for higher order, can *formally* be used at small scales where PT is not expected to hold. It was observed in scale invariant $N$-body simulations (Colombi et al. 1995; Colombi et al. 1996) and observations (Szapudi et al. 1996), that this formal procedure gives an excellent fit for the higher order cumulants, even though the resulting $n_{\rm eff}$ is no longer the local slope of the power spectrum; rather, it is a formal parameter that proves to be extremely useful for characterizing data. In scale-invariant simulations it was found that a steepening occurs in terms of $n_{\rm eff}$; i.e., the distribution in terms of its cumulants at nonlinear scales is equivalent to another weakly nonlinear distribution but with a steeper power spectrum.

The solid lines in Figure 8 show the least-square fit for $n_{\rm eff}$ in all the large simulations. Up to sixth-order quantities were used, and the error bars were obtained formally by

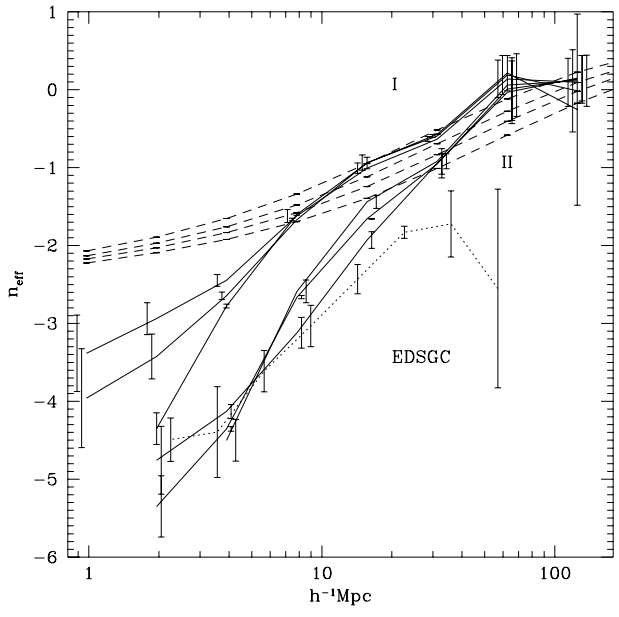


FIG. 8.—Measurements of the higher order moments summarized in terms of $n_{\rm eff}$ of EPT. The upper three solid lines with error bars display the fitted $n_{\rm eff}$ for simulations $I_a$, $I_b$, $I_c$, increasing in this order on scales of 2 $h^{-1}$ Mpc. The lower three solid lines correspond to $II_a$, $II_b$, $II_c$, decreasing in this order on scales of 15 $h^{-1}$ Mpc. The upper dashed line is the theoretical local slope of the power spectrum for the $\Omega = 1$ simulations; the lower three dashed lines are the same for $II_a$, $II_b$, $II_c$, decreasing in this order. The dotted lines with error bars show the measurements of $n_{\rm eff}$ in the EDSGC survey. The error bars on this figure were determined by calculating $n_{\rm eff}$ from $S_3$ alone and comparing it with a simultaneous fit using $S_3, \ldots, S_6$. The latter is displayed, and the difference of the two is an indication of the accuracy.

calculating $n_{\rm eff}$ from $S_3$ alone. This takes into account the inaccuracy of the higher order moments relative to the third-order moments, as well as the possible variance in EPT, an approximate phenomenological relation. This prescription, however, cannot account for any absolute errors on the measurements of the $S_N$'s. Figure 8, the most sensitive summary of the results of the paper, shows $n_{\rm eff}$ on a *linear* scale.

The upper three solid lines correspond to $I_a$, $I_b$, $I_c$ in increasing order on scales of $2\,h^{-1}$ Mpc, and the lower three correspond to $II_a$, $II_b$, $II_c$ in decreasing order on scales of $15\,h^{-1}$ Mpc. The $\Omega = 1$ and $\Omega < 1$ simulation groups are tightly together, while the two groups differ from each other. In the weakly nonlinear regime the agreement is excellent between PT theory and the measured $S_N$'s, since the $n_{\rm eff}$ is extremely close to the theoretical slope of the power spectrum: the upper dashed line shows the theoretical prediction for $\Omega = 1$, and the lower three dashed lines show the prediction for $\Omega < 1$ for $II_a$, $II_b$, $II_c$ in decreasing order. Note that the actual $\Omega$ dependence of the $S_N$'s, which is extremely small (Bouchet et al. 1992), was not taken into account for the theoretical prediction; simply the local slope of the power spectrum is plotted. Since the power spectrum is slightly different for the $\Omega < 1$ simulations, they behave differently in this regime. Because of nonlinearities at smaller scales, PT is not a good approximation; however, EPT still is. This can be seen from Figure 4, where triangles show the $S_N$'s formally corresponding to the fitted $n_{\rm eff}$. The agreement is excellent above $l_{50}$ on all figures, except perhaps for $II_c$ where it is only a good approximation above $2l_{50}$ for the higher orders. Thus $n_{\rm eff}$ at each scale is good representation of the data, providing a natural framework for comparison. In the highly nonlinear regime, a steepening compared to the PT value is present, which is apparent relative to the dashed line on the figure. This is very similar to the effect observed in scale invariant simulations by Colombi et al. 1995. Note also that as the simulations become more relaxed, EPT becomes more accurate. This suggests that the breakdown at small scales is caused only by inaccuracies introduced by dynamical discreteness effects at small scales. In Figure 8 the same effect shows up as a fanlike spreading of the curves, corresponding to a slight decrease of the $S_N$'s as a function of time, as discussed before. When particle discreteness is accounted for, the $S_N$'s appear to be approximately time-independent to a degree similar to the weakly nonlinear regime even at highly nonlinear scales. On the other hand, the difference between the two types of simulations is real, as it is observed at the reliable scales. Note that by construction the error bars of the figure cannot reflect systematics from particle discreteness.

The framework provided by $n_{\rm eff}$ is ideal for comparison with observations. The same type of calculation was performed by Szapudi et al. 1996 using the EDSGC survey. Their results agree well with the corresponding $S_N$'s from the Automatic Plate Measuring Facility Survey (Gaztañaga 1994; Szapudi & Gaztañaga 1998). The $n_{\rm eff}$ from the EDSGC is plotted with dotted lines. Although the split between the different output times could be artificial, as noted above, the difference between the $\Omega = 1$ and $\Omega < 1$ simulations is real. The comparison with the EDSGC data clearly favors the $\Omega < 1$ curves. Note a subtlety of the comparison shown here: the scales given with the deprojected $S_N$'s in Szapudi et al. 1996 are simply $D\theta$, where $D = 370\,h^{-1}$ Mpc is the depth of the catalog and $\theta$ is the angle of the sides of the square window used for counts in cell. Since the simulation uses cubical cells, the comparison with $D\theta$ is not appropriate. On the figure a simple approximation is used: the volume of the effective cone (or pyramid) is equated to the volume of the cells in the simulations. More precisely, $D^3\theta^2/3 = l^3$ was assumed, where $l$ is the side of the cubes in the simulations. If $\theta_{\rm deg}$ is expressed in degrees, the scale transformation is $l = 35.9\theta_{\rm deg}^{2/3}\ h^{-1}$ Mpc as opposed to the usual $l = 6.5\theta_{\rm deg}\ h^{-1}$ Mpc. Comparing the volumes should be a reasonable approximation on small scales, where virialization erases any configuration dependence, but it is expected to break down on larger, weakly nonlinear scales, where the elongated pyramids might have different $S_N$'s than the equivalent cubes (Scoccimarro et al. 1999).

While it is clear from the figure that the data favor the $\Omega < 1$ models, let us use a toy model of biasing to quantify this statement. This should be reasonably accurate in the weakly nonlinear regime, even though the configuration dependence of the higher order moments start to enter the picture. Here we use the *Ansatz* $S_3 = 34/7 - (n_{\rm eff} + 3)$ from EPT and the leading order bias formula $S_3^g = S_3/b + 3b_2/b^2$ (Fry & Gaztañaga 1993), where the galaxy field is expanded in a Taylor series as $\delta^g = b\delta + b_2\,\delta^2/2 + \ \ldots$ and the superscript $g$ signifies galaxies. If one formally applies EPT for the (possibly) biased galaxy field, as was done in the case of EDSGC (Szapudi et al. 1996), it is possible to express $b_2$ in terms of the measured effective indices of the galaxies $n_g$ as

$$b_2 = -(b/21)(13 - 13b - 7n_{\rm eff} + 7bn_g)\ . \qquad (7)$$

In this equation $b$ is fixed by the $\sigma_8$ of the simulation, and $n_{\rm eff}$ and $n_g$ are the measured effective index in the simulation and in the galaxy catalog, respectively. Note that the two-point functions of the different time outputs differ essentially only in the amplitude within the studied scale range; thus bias is approximately independent of scale and can be described by $\sigma_8$. The results of such a model are shown in Figure 9. The figure plots $b_2$ against scale from 4–20 $h^{-1}$ Mpc. Note that the two outlying points above 20 $h^{-1}$ Mpc in Figure 8 are caused by edge effects in the EDSGC survey (Szapudi et al. 1996). The curves in increasing order represent simulations $II_c$, $II_b$, $I_c$, $II_a$, $I_b$, $I_a$. The interpretation of the figure is not straightforward, since the leading order calculations are only expected to work on large scales, where the errors of the EDSGC measurements

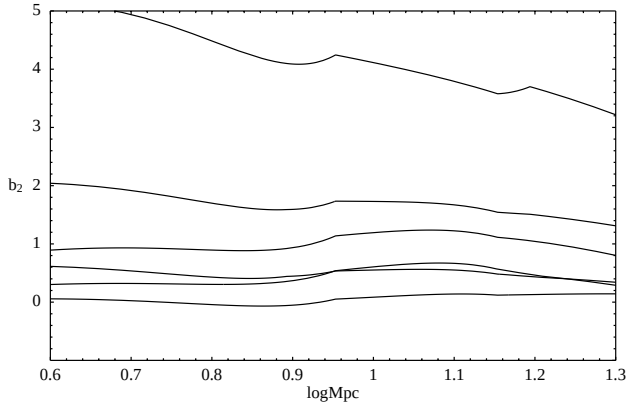


FIG. 9.—First nonlinear coefficient in the Taylor expansion of the bias, $b_2$, is displayed for the different simulations as a function of log scale. The curves, in increasing order, are for $II_c$, $II_b$, $I_c$, $II_a$, $I_b$, $I_a$.

and the configuration dependence are becoming increasingly influential. Nevertheless, it is intriguing that the $\Omega = 0.3$, $b = 1$ ($II_c$) model requires no nonlinear biasing in terms of $b_2$ within the errors, in contrast with all the other models. Thus it is possible to explain the higher order distribution of counts in cells in the EDSGC without invoking linear or nonlinear biasing. The minimal assumption that the EDSGC galaxies trace mass satisfies the above model, but none of the others examined. Occam's razor rejects them as they all need significant nonlinear biasing.

Note, however, that Figure 9 is not the ultimate answer. More investigations of the nonlinear bias are needed, where the above simple theory breaks down, because of the nonlinearities and stochasticity. Also, the configuration dependence can be modeled more accurately by the use of artificial catalogs with realistic selection functions, which employ pyramid shape cells as used in the EDSGC. In addition, more data (especially with redshifts) with larger dynamic range toward large scales will turn this argument into a more quantitative result. The possible extensions are left for future work, while the data requirements will be met by the new generation of galaxy surveys. In the near future, the SDSS, and the Two Degree Field Survey will determine the higher order moments with similar accuracy to the present simulations. Comparison of the future data with the results reported here will strongly constrain biasing models.

This paper presented the measurements of cumulants of counts in cells in CDM and ΩCDM simulations. These high-resolution simulations together with a pair of new measurement algorithms enabled us to explore a larger dynamic range with smaller errors than previously was possible. A careful attempt was made to determine the range of reliable scales, and a fully nonlinear theoretical error calculation was performed as well. It was found that, via PT and EPT, the results can be efficiently represented by the effective index, $n_{\rm eff}$. In the weakly nonlinear regime excellent agreement was found with PT, while at smaller scales a nonlinear plateau found in scale-invariant simulations was confirmed. At small scales, the agreement with EPT was found to be remarkably good. The CDM results are qualitatively similar to the scale invariant simulations in all respect considered. The time-dependency of the cumulants appears to be negligible at all scales, if particle discreteness is correctly taken into account. A comparison with observations revealed that the $\Omega < 1$ model is consistent with the higher order correlations of the EDSGC galaxies without the need of biasing. The rest of the models examined need substantial nonlinear biasing to be reconciled with the data. Finally, the error formulae of SC and SCB provide a good approximation for the errors on higher order statistics measured in $N$-body simulations.

I. S. would like to thank Stephane Colombi for stimulating discussions, Carlton Baugh and Shaun Cole for useful suggestions, and Albert Stebbins for instrumental help. I. S. was supported by the PPARC rolling grant for Extragalactic Astronomy and Cosmology at Durham. This work was also supported by DOE and NASA through grant NAG-5-2788 at Fermilab, and by a NASA HPCC-ESS grant. The simulations were carried out on the Pittsburgh Cray T3E, the Goddard Space Flight Center Cray T3E, and the Cray T3D at the Arctic Regional Supercomputing Center.

## APPENDIX

Here we discuss the algorithms used to calculate counts in cells. While from theoretical point of view it would be ideal to use the infinitely oversampling algorithm of Szapudi 1997 to estimate the distribution of counts in cells (SC), this would be unrealistically slow even with present day computers for 47 million particles in three dimensions. The algorithms discussed here provide an efficient way to sample with $\simeq 10^9 \times 10^{14}$ cells in less than 8 CPU hr on a typical workstation with approximately 1 GB of memory. The two methods are complimentary to each other, one for large and the other for small scales, with ample overlap between the two. They are outlined next.

### A1. LARGE SCALES

The algorithm for large scales will be explained in one dimension for simplicity. The generalization for arbitrary dimensions is obvious. The three dimensional version was used in the calculations of this paper.

The computations are performed on the largest possible grid with $N$ segments that can be fitted into the memory of the computer: this determines the smallest possible scale $L/N$, where $L$ is the box size and $N$ is the base sampling. A hierarchy of scales is used, with the scale at a given level being twice the scale at one level lower. The results one step lower in the hierarchy are used to keep the number of sampling cells constant even at the largest scales. Counts in cells can be straightforwardly calculated on the resolution scale of the grid, i.e., the smallest scale considered. For the calculation at twice the previous scale the sum of two cells are always stored in one of the cells, for instance, in the one with smaller index. Because of the periodic boundary conditions, auxiliary storage is required to calculate the sum of the values in the rightmost cell (if the summations was done left to right), since its right neighbor is the leftmost cell, which was overwritten in the first step. After these preparatory steps counts in cells can again be calculated from the $N$ numbers representing partially overlapping cells. For the next level, twice the previous scale, one needs the sum of four original resolution cells, a calculation simply done by summing every other cell of the previous results into one cell. At this level, two auxiliary storage spaces are needed because of the periodicity. In general, at each level in the hierarchy two cells of the previous results are summed as a preparatory step, and counts in cells are calculated simply by computing the frequency distribution of the $N$ sums stored in the main grid. Auxiliary storage is needed for those rightmost cells, which have the periodic neighbors on the left end.

In $D$ dimensions $2^D$ cells are summed in the preparatory step, and the auxiliary storage space enlarges the original hypercube. In our case the main grid was $512^3$, resulting in a $1\ h^{-1}$ Mpc spacing of $1.3 \times 10^8$ cells. Further precision could be achieved by oversampling the original grid, that is, shifting it by a fraction of a resolution cell. Our CPU resources allowed for

one independent shifting in each direction, which resulted in an 8 times oversampling of the original grid; that is, $C = 1.1 \times 10^9$ cells at each scale from $\simeq 1$ to $\simeq 128\ h^{-1}$ Mpc, i.e., a quarter of the length of the box.

### A2. SMALL SCALES

The above procedure is limited at small scales by the largest grid that will fit into the memory of the computer. Therefore an alternative technique was adapted for small scales using the original oct-tree data-structure of tree $N$-body codes. This is an efficient representation of a sparse array, since at small scales most of the cells are empty in a grid spanning the simulation. The tree is built up recursively, by always dividing the particles into two groups based on which half of the volume they belong to. The same function is called on both halves with the corresponding particles until there is no particle in the volume or the scale becomes smaller than a predetermined value. At each level the scale and the number of particles are known, and when an empty volume is reached, all contained volumes are also empty. These two observations are enough to insert the bookkeeping needed to calculate counts in cells at all scales while the tree is built. The number of sampling cells at each level are $2^l$, where $l$ is the level; the original box is represented by $l = 0$. Toward smaller scales the number of cells increases. When $N^3 = 2^l$, where $N$ is the size of the largest grid of the previous algorithm, the two techniques should (and do) give the exact same answers. At larger scales the previous algorithm is superior, since $N > 2^l$, whereas this algorithm becomes useful at smaller scales. Just as above, this procedure can be further improved by shifting the particles slightly before calculating the tree. However, since this hierarchy of grids has different numbers of cells, random shifts are more advantageous. Shifting by a fraction of the smallest scale would not exhaust the possibilities for any larger scale, whereas shifting by a fraction of the largest grid might not shift the underlying grids at all. With the introduction of random shifts (oversampling grids), the dynamic range of the two algorithms develops a substantial overlap, which is useful for testing. According to Figure 3, the algorithms produced essentially the same higher order moments in the overlap range of five powers of two.

### REFERENCES


Baugh, C. M., Gaztañaga, E., & Efstathiou, G. 1995, MNRAS, 274, 1049
Bernardeau, F. 1992, ApJ, 292, 1
———. 1994, ApJ, 433, 1
———. 1995, A&A, 301, 309
Bouchet, F. R., & Hernquist L. 1992, ApJ, 400, 25
Bouchet, F. R., Juszkiewicz, R., Colombi, S., & Pellat, R. 1992, ApJ, 394, L5
Bouchet, F. R., Schaeffer, R., & Davis, M. 1991, ApJ, 383, 19
Bouchet, F. R., Strauss, M. A., Davis, M., Fisher, K. B., Yahil, A., & Huchra, J. P. 1993, ApJ, 417, 36
Colombi, S., Bernardeau, F., Bouchet, F. R., & Hernquist, L. 1997, MNRAS, 287, 241
Colombi, S., Bouchet, F. R., & Hernquist, L. 1996, ApJ, 465, 14
Colombi, S., Szapudi, I., & the Virgo Collaboration. 1999, in Proc. ESO/MPA conf., Evolution of Large Scale Structure, ed. A. J. Banday et al. (Garching: Twin Press), in preparation (astro-ph/9810179)
Frieman, J. A., & Gaztañaga, E. 1999, in preparation
Fry, J. N. 1994, Phys. Rev. Lett., 73, 215
Fry, J. N., & Gaztañaga, E. 1993, ApJ, 413, 447
Fry, J. N., & Peebles, P. J. E. 1978, ApJ, 221, 19
Gaztañaga, E. 1992, ApJ, 319, L17
———. 1994, MNRAS, 268, 913
Gaztañaga, E., & Baugh, C. M. 1995, MNRAS, 273, L1
Governato, F., Babul, A., Quinn, T., Tozzi, P., Baugh, C. M., Katz, N., & Lake, G. 1999, MNRAS, submitted
Gunn, J. E., & Knapp, G. R. 1993, in Sky Surveys: Protostars to Protogalaxies, G. Neugebauer, ed. (San Francisco: ASP), 267
Jain, B. 1997, MNRAS, 287, 687
Juszkiewicz, R., Bouchet, F. R., & Colombi, S. 1993, ApJ, 412, L9
Kim, R. S., & Strauss, M. A. 1998, ApJ, 493, 39
Matarrese, S., Verde, L., & Heavens, A. F. 1997, MNRAS, 290, 651
Meiksin, A., Szapudi, I., & Szalay, A. 1992, ApJ, 394, 87
Peacock, J. A., & Dodds, S. J. 1996, MNRAS, 280, 19
Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Scoccimarro, R. 1998, MNRAS, 299, 1097
Scoccimarro, R., Colombi, S., Fry, J. N., Frieman, J. A., Hivon, E., & Mellott, A. 1998, ApJ, 496, 586
Scoccimarro, R., Szapudi, I., & Frieman, J. A. 1999, in preparation
Stadel, J., & Quinn, T. 1999, in preparation
Szapudi, I. 1997, ApJ, 497, 16
———. 1998, MNRAS, 300, L35
Szapudi, I., & Colombi, S. 1996, ApJ, 470, 131 (SC)
Szapudi, I., Colombi, S., & Bernardeau, F. 1999, in preparation (SCB)
Szapudi, I., Dalton, G., Efstathiou, G. P., & Szalay, A. 1995, ApJ, 444, 520
Szapudi, I., & Gaztañaga, E. 1998, MNRAS, 300, 493
Szapudi, I., Meiksin, A., & Nichol, R. C. 1996, ApJ, 473, 15
Szapudi, I., & Szalay, A. 1993, ApJ, 408, 43
———. 1997a, ApJ, 481, L1
———. 1998, ApJ, 494, L41
Szapudi, I., Szalay, A., & Boschán, P. 1992, ApJ, 390, 350