THE ASTROPHYSICAL JOURNAL, 479:606–615, 1997 April 20
© 1997. The American Astronomical Society. All rights reserved. Printed in U.S.A.

THE SMALL-SCALE VELOCITY DISPERSION OF GALAXIES: A COMPARISON OF COSMOLOGICAL SIMULATIONS

RACHEL S. SOMERVILLE,[1] JOEL PRIMACK,[1] AND RICHARD NOLTHENIUS[2]
Received 1996 April 9; accepted 1996 November 12

ABSTRACT

The velocity dispersion of galaxies on small scales ($r \sim 1\ h^{-1}$ Mpc), $\sigma_{12}(r)$, can be estimated from the anisotropy of the galaxy-galaxy correlation function in redshift space (Davis & Peebles). We apply this technique to "mock catalogs" extracted from N-body simulations of several different variants of cold dark matter–dominated cosmological models, including models with cold + hot dark matter, to obtain results that may be compared consistently with similar results from observations. We find a large variation in the value of $\sigma_{12}(1\ h^{-1}$ Mpc) in different regions of the same simulation. We investigate the effects of removing clusters from the simulations, using an automated cluster-removing routine, and find that this reduces the sky variance but also reduces the discrimination between models. However, studying σ_{12} as clusters with different internal velocity dispersions are removed leads to interesting information about the amount of power on cluster and subcluster scales. We compute the pairwise velocity dispersion directly, in order to check the Davis-Peebles method, and find agreement of better than 20% in all the models studied. We also calculate the mean streaming velocity and the pairwise peculiar velocity distribution in the simulations, and compare these with the models used in the Davis-Peebles method. We find that the model for the mean streaming velocity may be a substantial source of error in the calculation of σ_{12}.

Subject headings: cosmology: theory — galaxies: clusters: general — galaxies: distances and redshifts — large-scale structure of universe

1. INTRODUCTION

The velocity dispersion of galaxies on small scales ($r \sim 1\ h^{-1}$ Mpc), combined with cluster abundance data on intermediate scales and the *COBE* normalization and galaxy peculiar velocity information on large scales, provides a strong constraint on cosmological models by constraining the shape of the matter power spectrum. In this paper, we investigate a method introduced by Peebles (1976, 1980) and Davis & Peebles (1983, hereafter DP83), which uses the anisotropy of the redshift-space correlation function to determine the pairwise velocity dispersion on small scales. We shall refer to this method as the Davis-Peebles method.

The galaxy-galaxy correlation function $\xi(r)$ is one of the canonical statistics used in studying large-scale structure. A related statistic is the redshift-space correlation function, $\xi(r_p, \pi)$, which is a function of the components of the separation in redshift space perpendicular (r_p) and parallel (π) to the line of sight. If the correlation function is isotropic in real space, it will be anisotropic in redshift space because of the peculiar velocities of the galaxies. Hence, the degree of anisotropy of $\xi(r_p, \pi)$ is a measure of the moments of the peculiar velocity distribution.

The first moment of the pairwise velocity distribution, $\overline{v_{12}}(r)$, is proportional to $\Omega_0^{0.6}$ if the galaxies' trace mass and density fluctuations are in the linear regime, where Ω_0 is the density of matter in units of the critical density at the present epoch. The second moment, σ_{12}, is the velocity dispersion and measures the kinetic energy of the galaxy distribution. This quantity has been used in combination with the cosmic virial theorem to estimate Ω_0 (DP83), although recently some authors have presented arguments that this calculation is not only plagued by extremely large uncertainties from cosmic variance (Fisher et al. 1994b), but that some fundamental assumptions in the usual formulation of the cosmic virial theorem may also be incorrect (Bartlett & Blanchard 1995).

The first calculation of σ_{12} on a relatively large survey was done by DP83. They calculated σ_{12} for the CfA1 redshift survey, a survey containing 1840 redshifts covering 1.83 sr in the north Galactic hemisphere (Huchra et al. 1983). In another paper (Somerville, Davis, & Primack 1997, hereafter SDP), we present a reanalysis of their work, in which we show that σ_{12} for this sample is extremely sensitive to the way in which corrections for infall into the Virgo Cluster are applied, and that the value of σ_{12} for this survey is dominated by the clusters. The same calculation was done on the Southern Sky Redshift Survey (SSRS1) (SDP; Davis 1988) and on the *IRAS* survey (Fisher et al. 1994b). The sensitivity of σ_{12} to clusters and infall corrections was also pointed out by Mo, Jing, & Börner (1993), who analyzed the CfA1, SSRS1, *IRAS*, and CfA slice samples in a variety of ways. Recently, σ_{12} has also been calculated for the CfA2/SSRS2 survey (Marzke et al. 1995) and the Perseus-Pisces survey (Guzzo et al. 1995). These calculations have shown a range of values of $\sigma_{12}(1)$ from 272 km s^{-1} for SSRS2 to 769 km s^{-1} for Perseus-Pisces (see Table 1 of SDP for a summary).

The values of σ_{12} usually quoted for simulations are calculated by measuring the dispersion of the pairwise peculiar velocity field directly, using the full three-dimensional position and velocity information for the halos (Davis et al. 1985; Gelb & Bertschinger 1994; Klypin et al. 1993). In real redshift surveys, not only are there errors introduced by edge effects and selection effects, but σ_{12} is extracted by fitting a model to the correlation function in redshift space, $\xi(r_p, \pi)$. This quantity is quite noisy, especially for samples with small numbers of galaxies. In addition, the procedure

[1] Physics Department, University of California, Santa Cruz, CA 95064.
[2] UCO/Lick Observatory, University of California, Santa Cruz, CA 95064.

involves a number of assumptions. It is a reasonable question, therefore, whether or not the values from simulations may be meaningfully compared with the observational values. Zurek et al. (1994) applied the Davis-Peebles method to mock redshift surveys extracted from simulations of a standard cold dark matter model and found that this yields a rather large range of values for σ_{12}. So far, there has not been a comparison of σ_{12} calculated using the Davis-Peebles method on "observed" simulations of different cosmological models. In this paper, we investigate the robustness of σ_{12} by using mock redshift surveys extracted from several different cold dark matter–dominated models and investigate the ability of this statistic to discriminate between such models. We estimate the sky variance and cosmic variance of the statistic and identify sources of error in the Davis-Peebles method. We also investigate the effects of removing clusters from the samples.

2. COSMOLOGICAL MODELS AND SIMULATIONS

The models and simulations are described in detail in Klypin, Nolthenius, & Primack (1997, hereafter KNP). All of the models studied here have Gaussian initial fluctuations with a Harrison-Zeldovich scale invariant spectrum, density parameter $\Omega = 1$ and Hubble parameter $h = 0.5$ ($H_0 = 100\,h$ km s^{-1} Mpc^{-1}).

The models represented are two variants of cold dark matter (CDM) and one variant of cold + hot dark matter (CHDM). The "standard" CDM model assumes that the fraction of the mass in the universe made of baryons is $\Omega_b = 0.1$, with the rest of the mass made up of a species of nonrelativistic, dissipationless particle (the cold dark matter). There are different ways of normalizing the spectrum: "unbiased" ($b = 1$), which assumes that the galaxy fluctuations trace the dark matter fluctuations, and "biased," which assumes that perhaps only especially large amplitude fluctuations lead to galaxy formation, i.e., $(\delta\rho/\rho)_{\rm galaxy} = b(\delta\rho/\rho)_{\rm mass}$. We have analyzed an "unbiased" ($b = 1$) simulation (which we call CDM1) and a $b = 1.5$ simulation (CDM1.5). In linear theory, this corresponds to rms fluctuations of mass in a sphere of 8 h^{-1} Mpc radius $\sigma_8 = 1$ for the unbiased model, $\sigma_8 = 0.667$ for the biased model, or a quadrupole in the angular fluctuations of the cosmic microwave background of $Q_{\rm ps-norm} = 12.8$ or 8.5 μK, respectively (KNP). The value of $Q_{\rm ps-norm}$ for the unbiased model was consistent with the lowest normalization quoted in the first-year *COBE* results (Smoot et al. 1992). It is too low in view of the more recent *COBE* results, which give $Q_{\rm ps-norm} = 18$ μK for $n = 1$ (Bennett et al. 1996), but for the purposes of the present paper, this should not pose a serious problem. The biased model (CDM1.5) is not compatible with the *COBE* measurement of $Q_{\rm ps-norm}$; however, it does predict approximately the correct number density of clusters (Frenk et al. 1990) and is useful for comparison.

The CHDM model assumes that the nonbaryonic matter is made of a cold component, as before, and a hot component, generally thought to be a massive neutrino. In "standard" CHDM, the ratio of cold to hot to baryonic matter is $\Omega_c/\Omega_\nu/\Omega_b = 0.6/0.3/0.1$, corresponding to a single massive neutrino with a mass of about 7 eV. The power spectrum in our model is normalized to $\sigma_8 = 0.667$, corresponding to $b = 1.5$ or $Q_{\rm ps-norm} = 17$ μK. More recent data and analyses favoring a smaller fraction of the mass in neutrinos, $\Omega_\nu \simeq 0.2$, are summarized in Primack et al. (1995) and Liddle et al. (1996). Our preliminary analysis of new simulations with $\Omega_\nu = 0.2$ indicates that σ_{12} in this model does not differ substantially from the $\Omega_\nu = 0.3$ model on the scales of interest in this paper.

The simulations were done using a standard particle-mesh (PM) code with a 512^3 force mesh. Each simulation has 256^3 cold particles, and the CHDM simulations have two additional sets of 256^3 hot particles with random thermal velocities corresponding to the Fermi-Dirac distribution (KNP). The size of the computational box is 50 h^{-1} Mpc, and the smallest resolved comoving scale is ~ 100 h^{-1} kpc.

The initial fluctuations were generated with the same random numbers for CHDM1 and both CDM simulations. After running the CHDM1 simulation, it was discovered that there were two mistakes in the initial conditions (see KNP): the fitting formula for the cold spectrum was too small, and the velocities were too large, both by about 20% on small scales. However, these effects are in phase and largely cancel. A new simulation with the same cosmological parameters but different initial conditions (and mistakes corrected) was run—we call this CHDM2. Extensive comparison of the two simulations has shown that power and velocity differences on small scales are no more than 5% at $z < 7$. In addition, there was a statistical fluke (of about 10% probability) in the initial conditions for both CDM simulations and CHDM1: the amplitude of the longest waves was a factor of 1.3–1.4 larger than that expected for the ensemble average, so the power on large scales is approximately a factor of 2 larger than typical. However, this could be considered a compensation for the finite size of the box or, in the case of CDM1, as compensation for the low normalization compared with *COBE*. Also, comparing CfA1 with the much larger APM survey shows that the CfA1 region has unusually high power at these scales (Nolthenius, Klypin, & Primack 1997, hereafter NKP; Baugh & Efstathiou 1994). In any case, a comparison of CHDM1 and CHDM2, which has a typical amount of power, gives us a measure of cosmic variance.

In order to compare our simulations with observations, we need to identify objects that correspond to galaxies. One approach is to assume that galaxies form in regions where the dark matter has collapsed to a sufficiently high density. However, a well-known problem with this procedure is that when the dark matter halos merge, they quickly lose all discernible substructure. This results in a halo mass distribution that includes a few very large mass ($\sim 10^{13}$–10^{15} $M_\odot$) halos in the final epoch of the simulations, which should probably be associated with the dark matter in cluster cores rather than with the halos surrounding single galaxies (Katz & White 1993; Gelb & Bertschinger 1994). Whether or not it is possible to resolve individual halos in the limit of infinite resolution depends sensitively on the inner density profiles of the halos (Moore, Katz, & Lake 1996). If the inner density profiles of the halos are isothermal or steeper (as suggested by the simulations of Carlberg 1994, Warren et al. 1992, and Crone, Evrard, & Richstone 1994), then the halos will survive in the limit of infinite resolution, although in practice the computational cost would be prohibitive. If the profiles are shallower than isothermal in the centers, as found by Katz & White (1993), Navarro, Frenk, & White (1996), and Xu (1996), then overmerging will occur, even in the limit of infinite resolution. In this case, it appears that there is no unique way to identify

galaxies in dissipationless simulations (Summers, Davis, & Evrard 1995).

In simulations with hydrodynamics and gas, the baryons can lose energy and form smaller clumps that remain distinct within the overmerged dark matter halos (cf. Evrard, Summers, & Davis 1994; Hernquist & Katz 1989; Cen & Ostriker 1992). This is probably what happens in the real universe—the dark matter halos surrounding individual galaxies are stripped when the galaxies fall into a larger potential well, resulting in a large dark matter halo containing many galaxies (a group or cluster). However, associating the baryon clumps in the current generation of simulations with galaxies is likely to be misleading since the details of galaxy formation depend on many interdependent processes, including dissipation, gasdynamics, star formation, and energy feedback from supernova (see Steinmetz 1996). Exactly how these processes affect galaxy formation is not well understood, and it is not possible to include these effects in N-body simulations on cosmological scales with the present computing capabilities. Clearly, for the present study of large-scale structure, it is necessary to attempt to make use of information from dissipationless simulations as best we can. Therefore, we have adopted a scheme, informed by results from simulations with gas and hydrodynamics (Evrard et al. 1994), for "breaking up" the overmerged dark matter halos. Although this procedure is ad hoc and has many limitations, it does go a step further toward enabling a realistic comparison between real redshift surveys and simulations than many previous analyses of simulations. In future papers, we will investigate the results of a more physically motivated scheme based on the approach of Kauffmann, Nusser, & Steinmetz (1995).

The procedure for breaking up the halos, assigning luminosities, and forming magnitude-limited "sky catalogs" that mimic the CfA1 redshift survey is described in detail in NKP. First, halos were identified as mesh cells with a sufficiently high, dark particle mass overdensity $\delta\rho/\rho$ at the end of the simulation ($z = 0$). Halos with a mass above a certain cutoff were broken up, and the fragments were assigned masses according to a Schechter distribution. Fragment velocities were chosen randomly from a Gaussian distribution using the rms velocity of the nearest neighbors as the dispersion. Luminosities were chosen randomly from a Schechter distribution with the same parameters as the CfA1 catalog and assigned to the halo fragments, assuming that higher luminosity corresponded to higher 1 cell mass. This results in a distribution with the same selection function as the CfA1 catalog. The broken-up dark matter halos are hereafter somewhat metaphorically referred to as "galaxies."

To construct the mock CfA1 catalogs, six different "home galaxies" were selected for each simulation, in such a way as to mimic the conditions of the CfA1 redshift survey: the home galaxies were required to lie in an area with local galaxy density within a factor of 1.5 of the local CfA1 density, and to have a Virgo-sized cluster about 20 Mpc away. An "observer" was placed on each home galaxy, and radial velocities along the line of sight were calculated for each galaxy. Catalogs were created with the same angular boundaries as the CfA1 north survey ($b > 40°$, $\delta > 0°$, 1.83 sr), and also all-sky catalogs containing all galaxies with $|b| > 10°$ (10.4 sr).

In order to check the effectiveness of our breakup procedure, we have verified that the galaxy-galaxy correlation function in real space follows a power law all the way down to the resolution limit. Without breakup, the correlation function falls below the power law on small scales. We have also checked that hierarchical scaling holds for the galaxies, i.e., that the volume-averaged three-point function is proportional to the averaged two-point function squared. We have found that the reduced skewness $S_3 = \bar{\xi}_3(r)/\bar{\xi}_2(r)^2$ is constant to a good approximation from $r = 0.5$ Mpc to $r = 20$ Mpc. These results suggest that our breakup procedure leads to a halo distribution with the expected clustering properties, at least to this order on these scales.

3. METHOD

In this section, we describe briefly the method used to extract the pairwise velocity dispersion σ_{12} from the redshift-space correlation function $\xi(r_p, \pi)$. Readers should refer to DP83 and Fisher et al. (1994a, 1994b) for more details of the method.

The correlation function in redshift space, $\xi(r_p, \pi)$, is estimated by counting the number of pairs in a bin in r_p (separation perpendicular to the line of sight) and π (separation parallel to the line of sight). It is normalized by constructing a catalog of Poisson-distributed points with the same selection function and angular limits as the data, and counting pairs between the data and the Poisson catalog:

$$1 + \xi(r_p, \pi) = \frac{n_R}{n_D}\frac{DD(r_p, \pi)}{DR(r_p, \pi)}, \tag{1}$$

where DD is the number of pairs between data and data, and DR is the number of pairs between the data catalog and a Poisson catalog. The quantities n_R and n_D are the minimum variance-weighted densities (see Davis & Huchra 1982) of the Poisson and data catalogs, respectively. In practice, we use a large ensemble of Poisson catalogs in order to reduce shot noise and to ensure that no bin has zero pair count.

Let $F(w|r)$ be the distribution function of velocity differences w for pairs of galaxies with vector separation r, and let $f(w_3|r)$ be the velocity distribution function averaged over the directions perpendicular to the line of sight. The first moment of $F(w|r)$ is $\overline{v_{12}}(r) \equiv -\langle[v(x) - v(x + r)] \cdot \hat{r}\rangle$, where the average is number (not volume) weighted. This quantity is also known as the mean streaming velocity, and assuming isotropy, it is a function only of the magnitude of r. Correspondingly, the first moment of $f(w_3|r)$ is $\langle w_3\rangle = y\overline{v_{12}}(r)/r$, where y is the component of r along the line of sight.

A reasonable form for the distribution function $f(w_3|r)$, parameterized by its moments, is

$$f(w_3|r) \propto \exp\left\{v\left[\frac{w_3(r) - \langle w_3(r)\rangle}{\sigma_{12}(r)}\right]^n\right\}. \tag{2}$$

It has been found empirically from studying observations and N-body simulations (Peebles 1976; Fisher et al. 1994b; Marzke et al. 1995) that on small scales, an exponential form ($n = 1$) fits the data better than a Gaussian ($n = 2$) or any higher power of the argument. We have also tested this assumption for our N-body simulations and find excellent agreement with an exponential in all the models up to scales of $r \sim 5\ h^{-1}$ Mpc. Recently, Sheth (1996) gave a derivation of the exponential form for the distribution function using the Press-Schechter approach. Adopting this form for

$f(w_3 \mid r)$, and using $r^2 = r_p^2 + y^2$ and $w_3 = \pi - y$, we have

$$1 + \xi(r_p, \pi) = \frac{H_0}{\sqrt{2}} \int \frac{dy}{\sigma_{12}(r)} [1 + \xi(r)] \times \exp \left\{ \frac{-\sqrt{2}}{\sigma_{12}(r)} \left| \pi - H_0 y \left[1 - \frac{\overline{v_{12}}(r)}{H_0 r} \right] \right| \right\}. \quad (3)$$

An approximation based on self-similar solutions of the BBGKY hierarchy leads to a form for $\overline{v_{12}}(r)$ (Davis & Peebles 1977):

$$\overline{v_{12}}(r) = \frac{F H_0 r}{[1 + (r/r_0)^2]} ; \quad (4)$$

F is an adjustable parameter of the model. The assumption of stable clustering (i.e., that the collapse of the cluster is balanced exactly by the Hubble flow) corresponds to $F = 1$. The results presented here are for $F = 1$ unless stated otherwise. Later on in this paper, we investigate this model, including the validity of the stable clustering assumption, by calculating $\overline{v_{12}}(r)$ in the N-body simulations. The velocity dispersion σ_{12} is then obtained by fitting the model of equation (3) to $\xi(r_p, \pi)$ estimated from equation (1).

4. RESULTS

One approach to comparing observations with simulations is to try to translate the space of observed quantities (i.e., redshift) into the space of the simulations (real space). Reflecting this approach, DP83 attempted to correct the measured redshifts by modeling the flow field around Virgo, which is in the foreground of the CfA1 survey. However, we have shown that the value of σ_{12} depends sensitively on the details of these corrections (SDP). Another approach is to "observe" the simulations in order to form simulated redshift surveys, which attempt to mimic as closely as possible what astronomers would actually observe if they lived in the universe of that simulation, and to then analyze the simulated and observed catalogs in exactly the same way. Taking this approach, we have selected catalogs from the simulations, requiring that a Virgo-sized cluster appear in the foreground as described in § 2, and do not include corrections for cluster infall in any part of the procedure.

Contour plots of $\xi(r_p, \pi)$ for the 10.4 sr sky catalogs are shown in Figure 1. For each simulation, $\xi(r_p, \pi)$ has been averaged over six views and smoothed. [For purposes of the contour plots only, $\xi(r_p, \pi)$ was calculated with linear bins in r_p and π. For the rest of the calculation, logarithmic bins

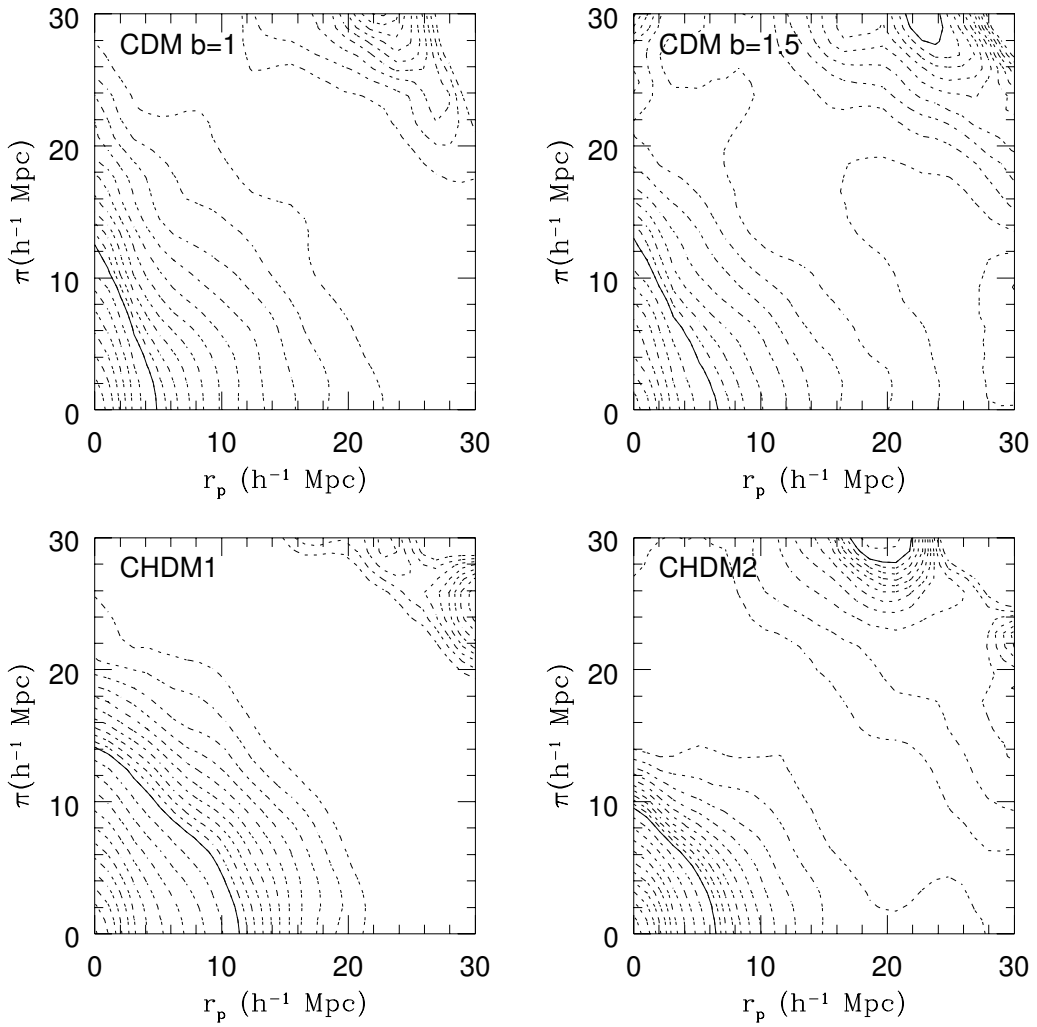

FIG. 1.—The contours show $\xi(r_p\ \pi)$ averaged over six views and smoothed. The contour spacings are $\Delta\xi = 0.1$ for $\xi < 1$ and $\Delta \log \xi = 0.1$ for $\xi > 1$, with the solid contour indicating $\xi = 1$.

were used in r_p, and linear bins in π]. There is a visible difference in the shape of the contours for the different models. In the absence of peculiar velocities, the contours would be perfectly circular segments. However, the velocity dispersion of clusters causes structures to appear elongated along the line of sight for small r_p (the familiar "fingers of God"), and infall causes the contours to appear compressed in the line-of-sight direction for larger r_p. This effect is much more pronounced in the CDM models, as expected, because of the larger velocity dispersions and infall velocities.

The values of r_0 and γ obtained from the inversion of $w(r_p)$ (see DP83) and used in the fit for σ_{12} are given in Table 3 below. The results for $\sigma_{12}(r)$ are shown in Figure 2. The points are the average of the results for the six views, and the error bars are standard deviations for the different views. These error bars reflect the "sky variance," or variation within the same simulation when viewed from different points, which, for this relatively small box size, is likely to be an underestimate of the true cosmic variance. The difference between CHDM1 and CHDM2, about 200 km s^{-1} at 1 h^{-1} Mpc, may give a better estimate of the possible cosmic variance. This is consistent with the large variance seen in recent calculations of σ_{12} in different redshift surveys (see SDP).

4.1. *Removing Clusters*

We attempted to address this apparent nonrobustness of σ_{12} by developing an algorithm to automatically remove clusters from catalogs in a way which could be applied consistently to both real and simulated data. It is no surprise that rich clusters are a major source of sample-to-sample variation in σ_{12}. Because $\xi(r_p, \pi)$ is a pair-weighted statistic, clusters tend to dominate it. It has been demonstrated in several papers (Mo et al. 1993; Zurek et al. 1994; SDP) that removing a single dominant cluster from a sample reduces significantly σ_{12}. The SSRS1 survey does not contain any clusters as rich as the Virgo or Coma clusters found in CfA1, and σ_{12} is much smaller for SSRS1 [$\sigma_{12}(1\ h^{-1}$ Mpc) $\simeq$ 320 km s^{-1}] than for CfA1 [$\sigma_{12}(1\ h^{-1}$ Mpc) $\simeq$ 620 km s^{-1}] (SDP). Fisher et al. (1994b) found $\sigma_{12}(1)$ for the *IRAS* redshift survey to be only about 320 km s^{-1}, much lower than the result for CfA1. The *IRAS* survey is dominated by dusty spiral galaxies and undercounts cluster centers by about a factor of 2 relative to optically selected surveys. Marzke et al. (1995) also found that removing Abell clusters of richness class $R \geq 1$ reduces significantly σ_{12} in the CfA2/SSRS2 survey, and that σ_{12} changes more drastically in regions of the survey where there were many rich clusters to begin with. In view of this evidence, we thought it would be interesting to see what happens to σ_{12} when the clusters are removed from the simulations. We therefore developed a method to remove clusters that can be applied consistently to both simulations and observations.

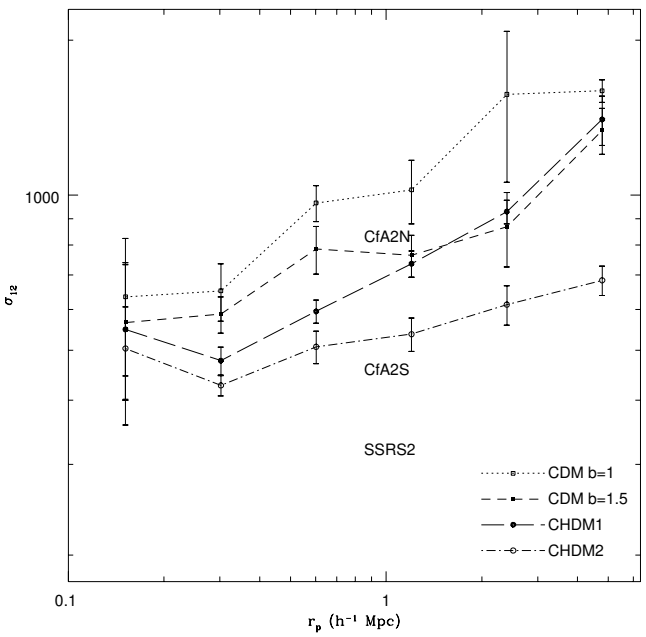

FIG. 2.—The velocity dispersion for the 10.4 sr sky catalogs, averaged over six views, with the standard deviation shown as error bars. The text markers CfA2N, CfA2S, and SSRS2 show the values from Marzke et al. (1995), for the CfA2/SSRS2 survey. They are shown offset to $r_p = 1$ for visibility but actually were calculated for exactly the same bin in r_p as our result.

Our algorithm is as follows. We divide the catalog into bins and identify bins with density fluctuations larger than a specified cutoff. We then calculate the luminosity-weighted centroid of all points lying within a cylinder of radius r_c and redshift interval $2h_c$ centered on the bin where the fluctuation was found. We use $h_c = 1000$ km s^{-1} and adjust r_c as described later. We then take a new cylinder centered on the centroid, calculate the new centroid, and continue to iterate in this way until the position of the centroid changes by less than some small value. Finally, we calculate the velocity dispersion of all the galaxies lying within the cylinder around the converged centroid and cut all the galaxies in this cylinder only if the velocity dispersion is greater than a cutoff σ_c. We have tested our algorithm by visualizing the surveys to make sure that the regions which are cut correspond to those that would be identified visually. The number of clusters identified depends on both parameters r_c and σ_c. We chose a value of r_c by plotting profiles of the clusters and identifying the radius at which the number density had dropped to the background level. The fiducial value we chose, $r_c = 2.0\ h^{-1}$ Mpc, is fairly close to the usual Abell radius. Table 1 shows the results of varying σ_c. As σ_c is lowered, the algorithm identifies more and more objects as "clusters"—of course, as we go to lower values of σ_c, we are really starting to identify objects that we would normally refer to as groups. We select $\sigma_c = 500$ km s^{-1} as corresponding to what are usually referred to as clusters because, at this cutoff, approximately 4%–9% of the galaxies are in clusters, which roughly agrees with observations. Table 2 shows properties of clusters identified at $\sigma_c = 500$ km s^{-1}.

The results for $\sigma_{12}(r)$ after cluster removal with $r_c = 2.0\ h^{-1}$ Mpc and $\sigma_c = 500$ km s^{-1} are shown in Figure 3. The sky variance has decreased somewhat for the CDM models but not significantly for the CHDM models. The cosmic variance between the two CHDM models has decreased slightly: $\sigma_{12}(1)$ for CHDM1 and CHDM2 now differs by about 150 km s^{-1} instead of 200 km s^{-1}. However, $\sigma_{12}(1)$ for the biased CDM model now lies between the two values for CHDM1 and CHDM2, and even standard CDM gives a value of $\sigma_{12}(1)$ that is within the view-to-view error bars of the CHDM1 value. Therefore, the small improvement in robustness acheived by removing the clusters appears to have been attained at the expense of discrimination between models. This problem cannot be solved by using a different

TABLE 1

PROPERTIES OF CLUSTERS

Simulation	σ_{cut} (km s^{-1})	N_{cl}	N_{gal}	f_{cl}	n_{cl} (10^{-7} Mpc^{-3})
CDM1	600	8	106	1.2	1.5
	500	37	773	8.8	6.9
	400	78	1605	18	14
	300	116	2254	26	22
CDM1.5	600	13	204	2.3	2.4
	500	45	776	8.7	8.3
	400	81	1398	16	15
	300	121	2142	24	22
CHDM1.....	600	5	89	1.0	0.93
	500	25	530	6.0	4.6
	400	78	1738	20	14
	300	130	2865	33	24
CHDM2.....	600	0	0	0	0
	500	15	368	4.3	2.8
	400	34	800	9.4	6.3
	300	82	1807	21	15

NOTE.—The properties were identified using the algorithm described in the text, averaged over six views, for the 10.4 sr sky catalogs. N_{cl} is the number of clusters with internal velocity dispersion greater than σ_{cut}, N_{gal} is the total number of galaxies in those clusters, f_{cl} is the percentage of galaxies in clusters, and n_{cl} is the number density of clusters.

value for σ_c, as can be seen from Figure 4. The discrimination becomes worse as we go to lower values of σ_c, and for higher values, there is very little change in the values of σ_{12}. The fact that $\sigma_{12}(1)$ appears to converge to almost the same value for all of the models as we lower σ_c is probably an indication that, once we remove the collapsed objects, σ_{12} is really a measure of Ω_0. In fact, originally this statistic was not designed as a tool to discriminate between models. It was hoped that it would be useful as a measure of Ω_0 (or of $\beta \equiv f(\Omega)/b$, to be precise, where b is the bias factor defined in § 2) (DP83; Davis 1995). It would be interesting to see what would happen if we remove clusters in simulations with different values of Ω_0.

The change in σ_{12} as a function of σ_c has a different shape for the different models. In Figure 5, we show the change in $\sigma_{12}(1)$ as a function of the number of clusters cut at different values of σ_c. It is interesting that the curve for CHDM1 lies on top of the curve for CHDM2 on this plot, even though the values of σ_{12} are very different. What this seems to indicate is that although the curve for CHDM1 extends farther to the right, indicating that CHDM1 has more clusters than CHDM2 (because of the excess large-scale power), the clusters themselves have the same velocity structure, because the small-scale power is the same. The CDM models have more power on small scales, and the clusters

TABLE 2

PROPERTIES OF CLUSTERS WITH $\sigma_{cut} \geq 500$ km s^{-1}

Simulation	$\bar{\sigma}_{cl}$ (km s^{-1})	$\bar{N}_{gal}$	$\sigma_{12}(1)$
CDM1	550 ± 49	20 ± 16	745 ± 114
CDM1.5	562 ± 45	17 ± 9	605 ± 38
CHDM1.....	552 ± 46	21 ± 10	652 ± 37
CHDM2.....	533 ± 23	25 ± 22	508 ± 41

NOTE.—The second column shows $\bar{\sigma}_{cl}$, the mean internal line-of-sight velocity dispersion of the clusters, with the standard deviation shown as the error. The third column shows the mean and standard deviation of the number of galaxies per cluster. The fourth column is $\sigma_{12}(1)$ after the clusters were removed.

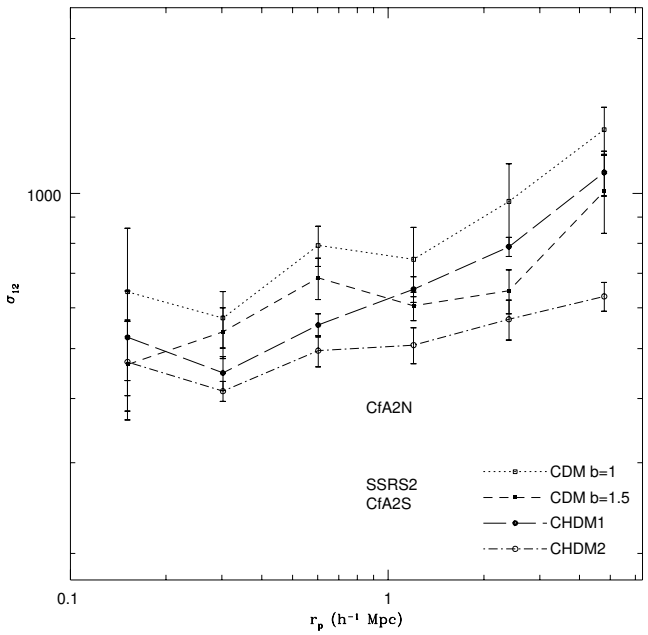

FIG. 3.—The velocity dispersion for the 10.4 sr sky catalogs with clusters removed, averaged over six views, with standard deviations shown as error bars. The text markers CfA2N, CfA2S, and SSRS2 show results from Marzke et al. (1995) for CfA2/SSRS2 with Abell clusters with richness $R \geq 1$ removed. See the comment in Fig. 2 on the horizontal offset.

have more kinetic energy, as indicated by the steeper rise in $\Delta\sigma_{12}(1)$. In addition, the curves extend farther to the right than CHDM2 because the CDM models also have a larger number density of clusters than a typical CHDM model. Plotted in this way, this quantity may give interesting information on the amount of power present on cluster and subcluster scales. It appears to be more robust and discriminatory than the actual value of $\sigma_{12}(1)$. However, it is not very useful at present because existing redshift surveys do not contain enough clusters to allow this quantity to be

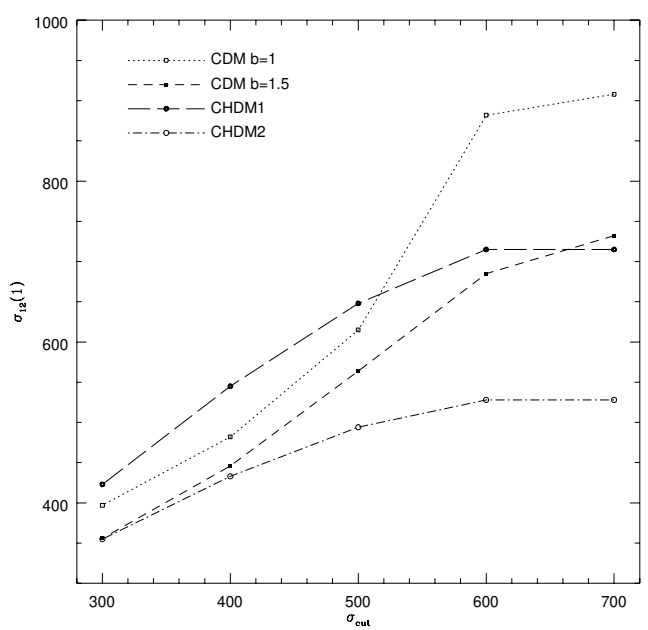

FIG. 4.—The velocity dispersion $\sigma_{12}(1)$, with clusters removed at different internal cluster velocity dispersion thresholds (σ_{cut}). These values are for a single mock catalog.

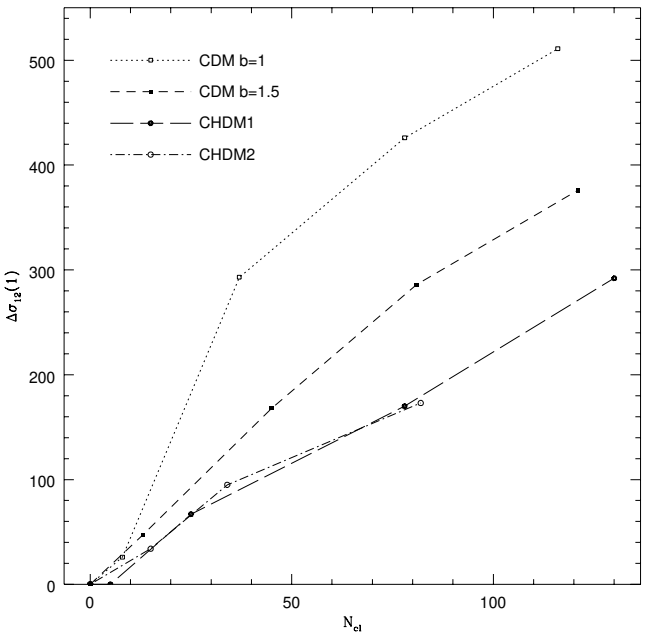

FIG. 5.—The change in velocity dispersion at $r = 1\ h^{-1}$ Mpc $[\Delta\sigma_{12}(1) = \sigma_{12}(1) - \sigma_{12}(1)_{\text{clusters cut}}]$ as clusters are removed at different thresholds, as a function of the number of clusters cut (N_{cl}). These results are for a single mock catalog.

calculated in the region where it is discriminatory. It would be interesting to study this in future large-volume surveys with good sky coverage.

4.2. *Form of Pairwise Peculiar Velocity Distribution and Mean Streaming*

A form for the pairwise peculiar velocity distribution must be assumed in the procedure we have used to calculate σ_{12}. As we discussed briefly, the form used by DP83 and other authors who have recently completed similar analyses is an isotropic exponential, which was proposed by Peebles (1976) because it appeared to fit observations. This functional form was further investigated by Fisher et al. (1994b) and Marzke et al. (1995) and found to be consistent with the *IRAS* and CFA2/SSRS2 redshift surveys, and with cold dark matter *N*-body simulations. However, we thought it would be interesting to study the form of the pairwise velocity distribution in real space for our different models in order to reevaluate whether or not this is the most appropriate form. We also take advantage of the fact that unlike in real observations, in the simulations, we know the real-space positions of the galaxies and their peculiar velocities separately. We compare the "true" value of σ_{12}, obtained by fitting directly to the peculiar velocity distribution, with the results of the Davis-Peebles procedure. We may then evaluate whether or not the "true" value of σ_{12}, uncomplicated by extracting it from the redshift space data, is a robust discriminator between models. In this way, we can test the accuracy of the Davis-Peebles method for extracting $\sigma_{12}(r)$.

We find that the exponential is an excellent fit to the distribution on small scales in all the models. On scales of $r \sim 5\ h^{-1}$ Mpc, the distribution begins to look a bit flatter at small w_3 than the exponential, and the distribution begins to approach a Gaussian at even larger separations. However, at the scales where this procedure is used ($\sim 1\ h^{-1}$ Mpc), using this model is not likely to be a significant source of error.

As we mentioned in § 3, following DP83 (and subsequent workers), we have modeled the mean streaming velocity $\overline{v_{12}}(r)$ by equation (4) in our fitting procedure. Marzke et al. (1995) and others have suggested that the observed scale dependence of σ_{12} may be merely an artifact of this term. We have investigated this quantity by computing it directly in the simulations. It is shown in Figure 6 along with the model for $F = 1$ and $F = 0.5$. The general form of the model holds on intermediate scales, but there is a large amount of variation between the different views, and the prediction of the model on scales of $1\ h^{-1}$ Mpc is in many cases quite inaccurate. The CDM models show a rather large negative mean streaming velocity on small scales, which may correspond to shell crossing. On average, the model with $F = 1$ appears to overpredict $\overline{v_{12}}$ for CHDM, while $F = 0.5$ may be a better fit. This is surprising because $F = 1$ corresponds to streaming that exactly cancels the Hubble flow on scales less than the correlation length (stable clustering); $F > 1$ corresponds to collapse on those scales, and $F < 1$ indicates that the clusters are not collapsing but, in fact, are expanding. This implies that stable clustering is not a good assumption for the CHDM models.

To show the dependence of the results on the assumed value of F, in Figure 7 we show the results obtained for σ_{12} (using the Davis-Peebles method on the "observed" sky catalogs, as before) when we do not include the streaming model in the fit, i.e., when we set $F = 0$. This systematically reduces the values of σ_{12} that we obtain and changes the scale dependence as expected. The CHDM models are now entirely flat over r_p. It is also possible to allow F to be fitted as a free parameter. The problem with this approach is that there is a degeneracy between F and σ_{12}—we tended to obtain lower values for both parameters when we allowed F to be fitted freely. In addition, having two free parameters increases the error on the fit. Apparently, the use of this model for the mean streaming, especially with the assumption of stable clustering ($F = 1$), could be a substantial source of error in the procedure.

Figure 8 shows the results obtained when σ_{12} was fitted directly to $f(w_3 \mid r)$ using the peculiar velocity information from the simulations. Although we use the real-space positions and peculiar velocities of the galaxies, we use the same galaxies as those in the magnitude-limited sky catalogs in order to facilitate a direct comparison of σ_{true} with the value obtained by the Davis-Peebles method (σ_{DP}). Table 3 shows σ_{DP} with and without the streaming model and σ_{true} at $1\ h^{-1}$ Mpc, averaged over the six views. The results for σ_{DP} are within 20% of σ_{true} for all the simulations. The agreement with σ_{true} is better for all the simulations, except CHDM2, if

TABLE 3

COMPARISON OF σ_{12}

Simulation	σ_{DP} ($F = 1$)	σ_{DP} (No Streaming, $F = 0$)	σ_{true}
CDM1.....	1024 ± 145	878 ± 133	862 ± 33
CDM1.5.....	764 ± 71	658 ± 62	677 ± 41
CHDM1.....	736 ± 43	607 ± 40	656 ± 52
CHDM2.....	537 ± 40	444 ± 35	543 ± 72

NOTE.—The comparison was computed using the method of Davis & Peebles with σ_{true}, computed by fitting directly to the pairwise peculiar velocity distribution.

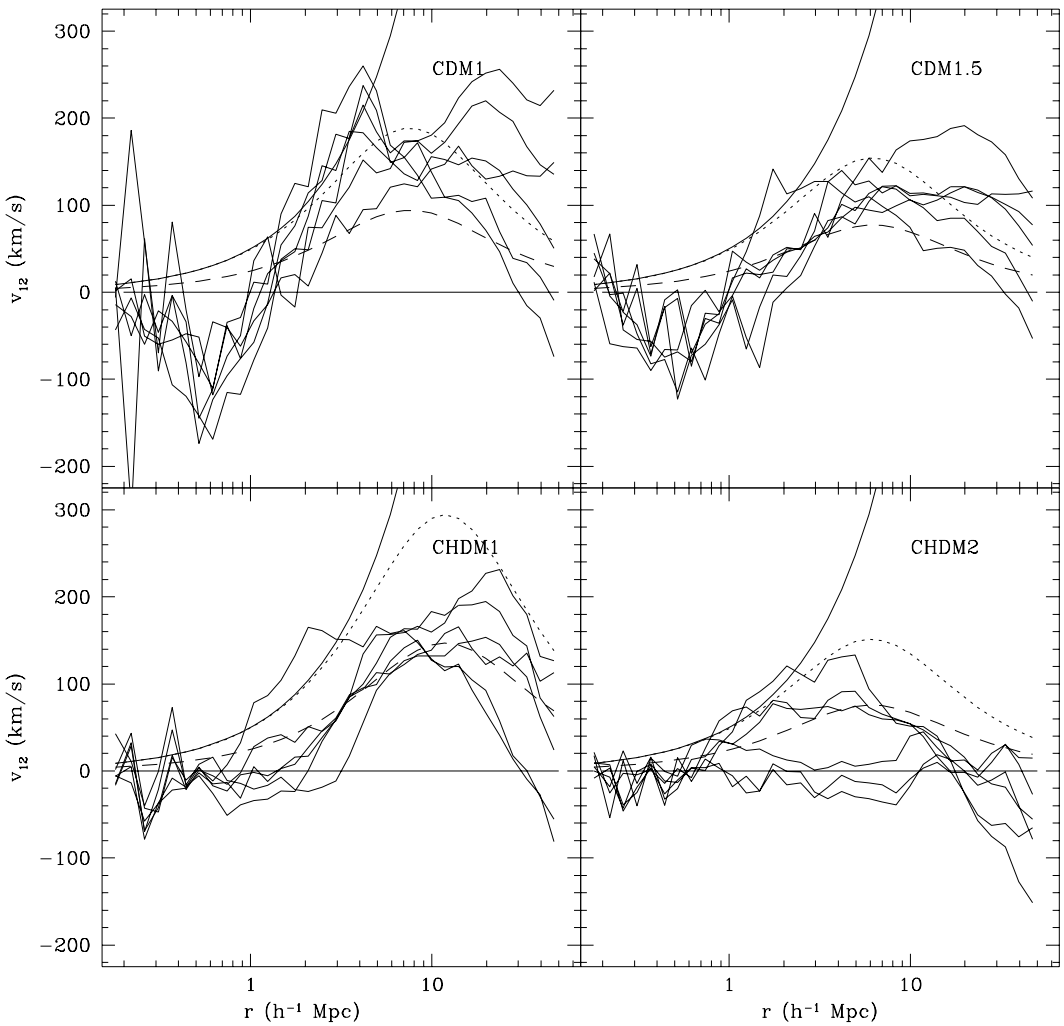

FIG. 6.—The first moment of the pairwise peculiar velocity distribution, or mean streaming $\overline{v_{12}}(r)$. The dotted line is the BBGKY model (eq. [4]) with $F = 1$, the dashed line is the model with $F = 0.5$, and the smooth solid line is the Hubble flow ($H_0 = 50$ km s^{-1} Mpc^{-1}). The results are shown for all six views in each simulation.

we do not include the streaming model ($F = 0$). The view-to-view variation is comparable for $\sigma_{\rm DP}$ and $\sigma_{\rm true}$, which suggests that the variation is not noise introduced by the Davis-Peebles method of extracting σ_{12}, but that it is intrinsic in the simulations and presumably in the real universe.

An important thing to keep in mind is that the "galaxies" (or halo fragments) were assigned velocities by hand (see § 2), and were assumed to be Gaussian and isotropic. However, skewness due to preferential infall is known to set in on scales of a few megaparsecs (Zurek et al. 1994). In addition, it is likely that real clusters (even the simulated clusters) are not really spherical and isotropic. It is unclear how big a role such effects may play in the results presented above, and this should be investigated further. However, this is beyond the scope of this paper.

5. DISCUSSION AND CONCLUSIONS

For the N-body simulations of three different cosmological models that we have analyzed, we find that the values of σ_{12} are considerably higher for CDM models than for CHDM models. The CHDM models, which give $\sigma_{12}(1) \sim 540$ (440) km s^{-1} for CHDM2 to 740 (600) km s^{-1} for CHDM1 are perhaps more consistent with the body of observational values taken as a whole. (Numbers in parentheses are for $F = 0$). Although $\sigma_{12}(1) = 647$ km s^{-1} for CfA2 north is marginally consistent, even with unbiased CDM [$\sigma_{12}(1) \sim 1024(880)$ km s^{-1}], $\sigma_{12}(1)$ for CfA2 south (367 km s^{-1}) and SSRS2 (272 km s^{-1}) are not (Marzke et al. 1995). However, because of the problems with galaxy identification discussed in § 2, we do not think that any model studied here should be ruled out on the basis of these results. Any existing large-volume N-body simulations would have similar problems. The point of this paper is precisely to suggest that it is premature to use σ_{12} to draw any strong conclusions about cosmological models. However, studying the simulations has given us other interesting information.

We have estimated the expected sky variance and cosmic variance of σ_{12} in our models. The following values are quoted for σ_{12} (1 h^{-1} Mpc). The sky variance (calculated as the standard deviation over six mock catalogs) ranges from $\sim$40 km s^{-1} for CHDM2 to $\sim$145 km s^{-1} for CDM1. The cosmic variance (between two simulations with different initial conditions) for CHDM is $\sim$200 km s^{-1}. The errors usually quoted for σ_{12} are formal errors on the fit (for which we obtain typically $\sim$40 km s^{-1}) and in general are underestimates of the actual statistical errors. We evaluate the accuracy of the Davis-Peebles method for extracting σ_{12}

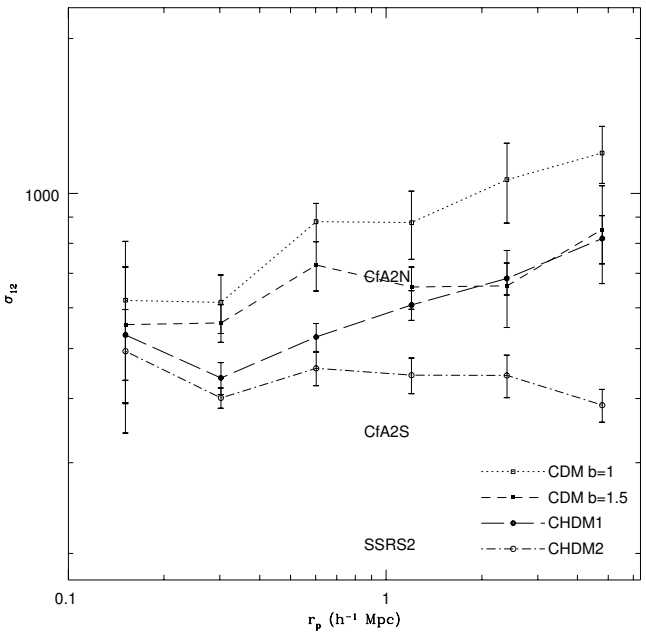

FIG. 7.—The velocity dispersion for the simulations with $F = 0$ used in the streaming model (eq. [4]), averaged over six views, with standard deviations shown as error bars. The text markers CfA2N, CfA2S, and SSRS2 show results with $F = 0$ from Marzke et al. (1995) for CfA2/SSRS. See the comment in Fig. 2 on the horizontal offset.

from redshift catalogs by comparing it with the results of computing the velocity dispersion directly in our mock catalogs. We obtain agreement of better than 20% for all of our models. This leads us to interpret the large range of values for σ_{12} obtained from different redshift surveys as an intrinsic variation that is due to the sensitivity of the statistic to the clusters contained in the sample, rather than being due to errors in the method.

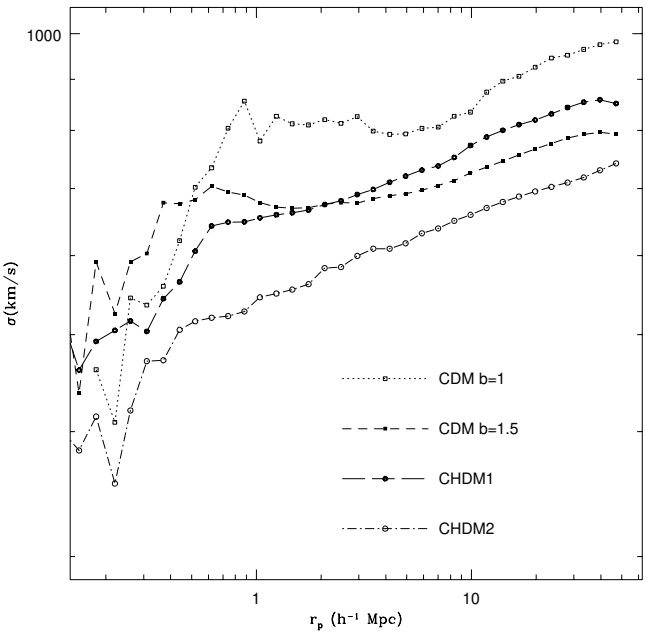

FIG. 8.—The second moment of the pairwise peculiar velocity distribution $f(w_3 \mid r)$, obtained by fitting the exponential model directly to $f(w_3 \mid r)$.

We have investigated the effects of removing clusters from the samples using an automated procedure. It was hoped that this might make σ_{12} a more robust statistic. However, we found that although this reduces the sample-to-sample variation in $\sigma_{12}(r)$ by a small amount, it actually reduces the ability of the statistic to discriminate between the cosmological models we studied. This may be due to the fact that all of our simulations are of $\Omega = 1$ models and that once clusters are removed, σ_{12} is really a measure of Ω_0. However, our study of the simulations suggests that the change in $\sigma_{12}(1)$ as a function of the number of clusters removed may be an interesting quantity to study in future redshift surveys.

We find that an exponential form for the pairwise peculiar velocity distribution is an excellent approximation on small scales ($r < 5\ h^{-1}$ Mpc) in all the models studied. Measuring the mean streaming $\overline{v_{12}}(r)$ directly from the simulations revealed that, although the general form of the model used in the Davis-Peebles method does hold, on small scales the measured values may deviate from it considerably. We found that stable clustering is a reasonable approximation in the unbiased ($b = 1$) CDM model but not in the CHDM models. The use of the BBGKY model for the mean streaming, especially with the assumption of stable clustering ($F = 1$), could be a substantial source of error in the Davis-Peebles method.

We have shown that σ_{12} is very sensitive to both the number and the properties of the clusters in a sample. This makes σ_{12} a poor constraint on cosmological models given the current situation with regards to both simulations and observations. Even the largest existing redshift surveys do not represent a fair sample of rich clusters. Also, current N-body simulations do not simulate clusters realistically because cluster properties are probably sensitive to non-gravitational physics such as gas hydrodynamics, star formation, and supernova feedback—effects that are impossible to include in large-volume simulations with current computing capabilities. As larger redshift surveys become available and it becomes possible to simulate clusters in a cosmologically relevant volume, perhaps the robustness of $\sigma_{12}(1)$ will improve. In fact, we have suggested a way in which the very sensitivity of σ_{12} to the properties of clusters could be used to define an interesting statistic for characterizing large-scale structure in larger samples. In the meantime, it is worthwhile to work on developing statistics that are discriminatory but less sensitive to the properties of clusters.

We thank M. Davis for useful discussions. R. S. S. is supported by a GAANN fellowship from the NSF. J. R. P. and R. N. are partially supported by grants from the NSF and NASA.

REFERENCES

Bartlett, J. G., & Blanchard, A. 1995, preprint
Baugh, C., & Efstathiou, G. 1994, MNRAS, 267, 32
Bennett, C. L., et al. 1996, preprint (http://xxx.lanl.gov/abs/astro-ph/9601067)
Carlberg, R. 1994, ApJ, 433, 468
Cen, R., & Ostriker, J. 1992, ApJ, 399, 331
Crone, M. M., Evrard, A., & Richstone, D. 1994, ApJ, 434, 402
Davis, M. 1988, in Cosmology and Particle Physics, ed. L. Z. Fang & A. Zee (New York: Gordon & Breach), 65
———. 1995, private communication

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
Davis, M., & Huchra, J. 1982, ApJ, 254, 437
Davis, M., & Peebles, P. J. E. 1977, ApJS, 34, 425
———. 1983, ApJ, 267, 465 (DP83)
Evrard, A. E., Summers, F. J., & Davis, M. 1994, ApJ, 422, 11
Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. P. 1994a, MNRAS, 266, 50
———. 1994b, MNRAS, 267, 927
Frenk, C. S., White, S. D. M., Efstathiou, G., & Davis, M. 1990, ApJ, 351, 10
Gelb, J. M., & Bertschinger, E. 1994, ApJ, 436, 491
Guzzo, L., Fisher, K. B., Strauss, M .S., Giovanelli, R., & Haynes, M. P. 1995, preprint
Hernquist, L., & Katz, N. 1989, ApJS, 70, 419
Huchra, J. P., Davis, M., Latham, D. W., & Tonry, J. 1983, ApJS, 52, 89
Katz, N., & White, S. D. M. 1993, ApJ, 412, 455
Kauffmann, G., Nusser, A., & Steinmetz, M. 1995, preprint (http://xxx.lanl.gov/abs/astro-ph/9512009)
Klypin, A., Holtzman, J., Primack, J., & Regos, E. 1993, ApJ, 416, 1
Klypin, A., Nolthenius, R., & Primack, J. 1997, ApJ, 474, 533 (KNP)
Liddle, A., Lyth, D. H., Shafi, R. K., & Viana, P. T. P. 1996, MNRAS, 282, 281
Marzke, R. O., Geller, M. J., da Costa, L. N., & Huchra, J. P. 1995, AJ, 110, 477
Mo, H. J., Jing, Y. P., & Börner, G. 1993, MNRAS, 264, 825
Moore, B., Katz, N., & Lake, G. 1996, ApJ, 457, 455
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Nolthenius, R., Klypin, A. A., & Primack, J. R. 1997, ApJ, in press (NKP)
Peebles, P. J. E. 1976, Ap&SS, 45, 3
———. 1980, The Large Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Primack, J. R., Holtzman, J., Klypin, A., & Caldwell, D. O. 1995, Phys. Rev. Lett., 74, 2160
Sheth, R. 1996, MNRAS, 279, 1310
Smoot, G. F., et al. 1992, ApJ, 396, L1
Somerville, R. S., Davis, M., & Primack, J. R. 1997, ApJ, in press (SDP)
Steinmetz, M. 1996, preprint (http://xxx.lanl.gov/abs/astro-ph/9608013)
Summers, F. J., Davis, M., & Evrard, A. 1995, ApJ, 454, 1
Warren, S. W., Quinn, P. W., Salmon, J. K., & Zurek, H. W. 1992, ApJ, 399, 405
Xu, G. 1996, private communication
Zurek, W., Quinn, P. J., Salmon, T. K., & Warren, M. S. 1994, ApJ, 431, 559