THE ASTROPHYSICAL JOURNAL, 468:62–74, 1996 September 1

THREE-DIMENSIONAL SHAPE STATISTICS: APPLICATIONS

SHAN LUO,[1] ETHAN T. VISHNIAC,[2] AND HUGO MARTEL[2]
Received 1995 December 8; accepted 1996 March 15

ABSTRACT

We apply the three-dimensional shape statistics developed in our previous paper to observations and simulations. We also develop an improved way of presenting the structures detected by the shape statistics over the one in the previous paper. We find significantly different clustering properties between the CfA survey and the Pisces-Perseus Supercluster survey on all scales we consider, implying that neither of them is a fair sample of the universe. Our results demonstrate that the shape statistics are powerful tools to detect clusters in a galaxy distribution, to distinguish filaments from pancakes, to discriminate different models of the large-scale structure, and to probe the dynamic evolution of structure formation.

Subject headings: cosmology: theory — galaxies: clusters: general — galaxies: structure — methods: statistical

1. INTRODUCTION

In a previous paper (Luo & Vishniac 1995, hereafter Paper I), we introduced the line and plane shape statistics to measure filaments and pancakes in a quantitative manner. The line shape statistic is defined as (a repeated index in a term means the index is summed over from 1 to 3)

$$S_l = \frac{1}{2\ \mathrm{Tr}\ (M)^3}\{-2\ \mathrm{Tr}\ (M)(M_{ij} - M_i M_j)(M_{ij} - M_i M_j) + \mathrm{Tr}\ (M)[\mathrm{Tr}\ (M) - M_i M_j]^2 + 6M_{ij}(M_{jk} - M_j M_k)(M_{ik} - M_i M_k) - 3[\mathrm{Tr}\ (M) - M_k M_k]M_{ij}(M_{ij} - M_i M_j)\}\ , \quad (1)$$

and the plane shape statistic is defined as

$$S_p = \frac{1}{\mathrm{Tr}\ (M)^3}\{4\ \mathrm{Tr}\ (M)(M_{ij} - M_i M_j)(M_{ij} - M_i M_j) - 4\ \mathrm{Tr}\ (M)[\mathrm{Tr}\ (M) - M_i M_i]^2 - 12M_{ij}(M_{jk} - M_j M_k)(M_{ik} - M_i M_k) + 12[\mathrm{Tr}\ (M) - M_k M_k]M_{ij}(M_{ij} - M_i M_j)\}\ , \quad (2)$$

where M_i and M_{ij} are the first and second moments of the point distribution in a window, and Tr (M) is the trace of M_{ij}. These shape statistics are constructed this way so that they vanish if the distribution in a window is spherically symmetric or constant except for a linear gradient across the volume. The quantity S_l (S_p) is unity for a line (plane) passing through the center of the window and zero for a plane (line) passing through the center of the window.

For a galaxy distribution, we center a window on each galaxy and compute the shape statistics, S, for the galaxies within this window. To calibrate the signal, we also calculate the values of the shape statistics, S_r, if the galaxies in the window obey a Poisson distribution. The signal of structure is then $S - S_r$. We take $\langle S - S_r\rangle$, the average of $S - S_r$ over all windows, as the final signal, and the standard deviation of $\langle S - S_r\rangle$ is used to compute the significance level of a nonzero $\langle S - S_r\rangle$.

Tests on toy models in Paper I have shown that these shape statistics are sensitive to the morphology of structures and the number and amplitude of filaments and pancakes in a distribution. These statistics can also detect structures from Poisson background points. In this paper, we apply these new statistics to data sets from observations and simulations. Our emphasis in this paper is still on the statistics themselves. That is, we are not going to construct a model that fits the current observations best, but instead we will study the behavior of the new statistics on more complicated data sets to see the effectiveness of these shape statistics for detecting differences between realistic data sets. Simulations allow us to study how the shape statistics measure the difference between models of structure formation, as well as their sensitivity to the evolution of gravitational clustering. We will also introduce *the map of structure orientation*, a new way of presenting the structures detected by these statistics. This method takes into account the relative strength of signals from different windows and gives the orientation of structure units.

In the next section, we introduce the map of structure orientation as a way of presenting the structure detected in a point distribution. We also compare this with the presentation we used in Paper I. In § 3, we apply the shape statistics to the CfA survey and the Pisces-Perseus Supercluster (PPS) survey. Our statistics show significant differences between these two samples, on all scales we consider. We also present the map of structure orientation for the samples, which visualizes the difference between the two samples detected by the shape statistics on different scales. In § 4, we apply the shape statistics to simulated galaxy distributions from four models of structure formation, namely, the cold dark matter (CDM) models with $\Omega_0 = 1$ and 0.5, and the hot dark matter (HDM) models with $\Omega_0 = 1$ and 0.5. We compare the behavior of the shape statistics for these models. We also study the sensitivity of the shape statistics for tracing the dynamic evolution of structure formation. Our results are summarized in § 5.

2. MAP OF STRUCTURE ORIENTATION

The presentation of the distribution of points given in Paper I is useful for revealing the distribution of clusters

[1] Astronomy Department, Cornell University, Ithaca, NY 14853; luo@spacenet.tn.cornell.edu.
[2] Astronomy Department, University of Texas at Austin, Austin, TX 78712; ethan@astro.as.utexas.edu, hugo@sagredo.as.utexas.edu.

and their members. It will be used to detect filaments and pancakes in a galaxy distribution. However, in terms of studying the shape statistics, it is not satisfactory in two respects. First, each point has equal weight, although the strengths of the signals from different windows can differ significantly. Second, the information of the orientation of a cluster in one single window is not revealed, although the global orientation of a well-defined filament or pancake can be inferred roughly from the relative positions of many points.

In that presentation, we plotted the centers of windows giving strong signals. Since each window is centered on a point, this has the advantage of giving real positions of points. However, it has the drawback that when a cluster is off center in a window, the signal is presented by a point away from the cluster. This will introduce extra scattering in the detected structures. To better trace the underlying structure, a presentation independent of the position of the window is desired. For this purpose, we replace the center of a window by the mean position of the points in the window. In this case, the positions are no longer real positions of points in a distribution.

We can measure the direction of a cluster in a window with the eigenvectors of the moment-of-inertia tensor of the distribution, $M_{ij} - M_i M_j$. For a filament, its direction (the major axis) is given by the eigenvector corresponding to the largest eigenvalue. For a pancake, the direction of its norm (the minor axis) is given by the eigenvector corresponding to the smallest eigenvalue.

In Figure 1 and Figure 2, we place a line segment at the mean position of the points in a window (the middle point of the line segment coincides with the mean position) if the signal $I = (S - \langle S_r \rangle)/\sigma_r > 2.5$, where S is the shape statistic from the data sample and $\langle S_r \rangle$ and σ_r are the mean and standard deviation of the shape statistic from calibrating Poisson samples. The line segment is pointed to the direction of the structure measured by the eigenvector described above. To reveal the relative strength of signals from different windows, we scale the length of the line segment according to the strength of the signal I. We call such plots the map of structure orientation.

The top panel, middle panel, and bottom panel in Figure 1 are for model A71 in Paper I with window diameters of 7.5, 20, and 40 units, respectively. In model A71, 10 pancakes, each containing 200 points and with axes 20, 20, and 2.5 units, are added to 8000 background points in a cube of size 100 units. The position and orientation of the pancakes are random. These panels are to be compared with Figures 8*c*–8*e* in Paper I, respectively. The top panel in Figure 1 shows that the eigenvector corresponding to the smallest eigenvalue of the moment-of-inertia tensor is a good measurement of the orientation of a pancake. The eigenvectors from windows covering a part of a pancake are perpendicular to the pancake. This can be seen most clearly from the edge-on pancakes.

Comparison of the middle panel in Figure 1 here and Figure 8*d* in Paper I shows that the mean positions of the points in a window have a smaller scatter than the centers of windows, as expected. Scaling a label by the strength of a signal also helps reveal structures better. In Figure 1, points away from a pancake have rather short length, as well as random orientations, indicating that they are field points, although they give signals above the threshold due to statistical fluctuations.

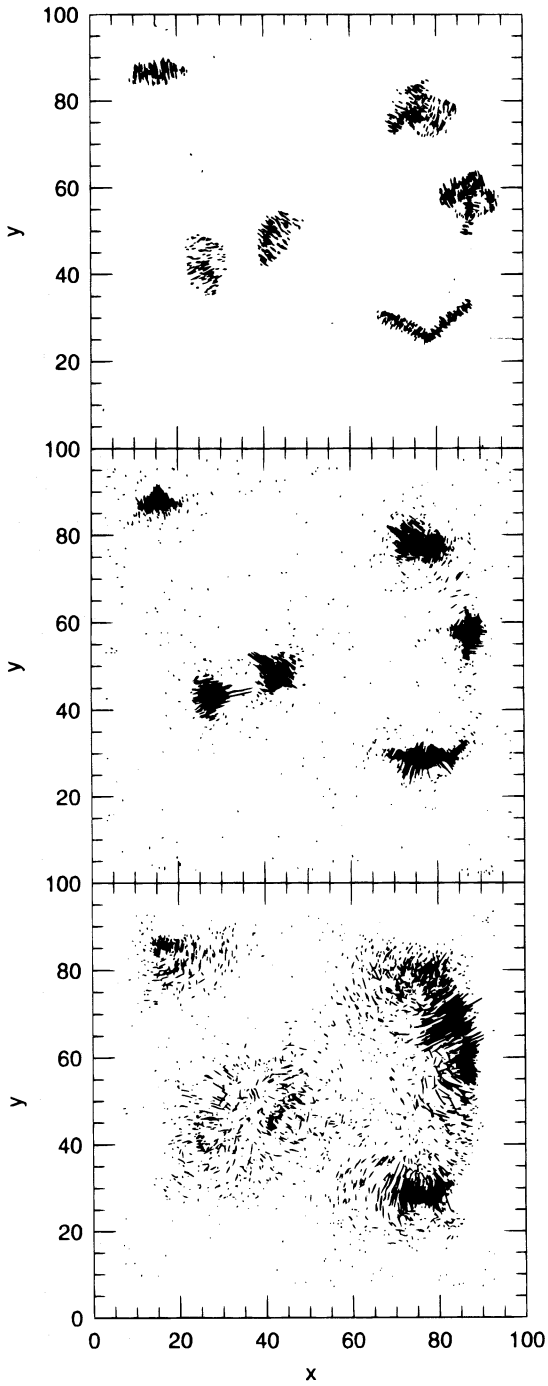

FIG. 1.—Directions of pancakes with a window diameter of 7.5 (*top panel*), 20 (*middle panel*), and 40 (*bottom panel*). The lengths of the line segments are proportional to the signal I_p.

The statistics become confused when the windows cover multiple structures, as shown in the bottom panel in Figure 1. Yet the coherence in the orientation of the line segments still clearly indicates the underlying structures.

The top, middle, and bottom panels in Figure 2 are for model A72 in Paper I with window diameters of 5, 20, and 25 units, respectively, and are to be compared to Figure 9*c*–9*e* in Paper I. In model A72, 10 spindles, each containing 50 points and with axes 20, 2.5, and 2.5 units, are added to 9500 background points in a cube of 100 units. The position and orientation of the spindles are random. We see a behavior similar to that for the case of pancakes: real structures

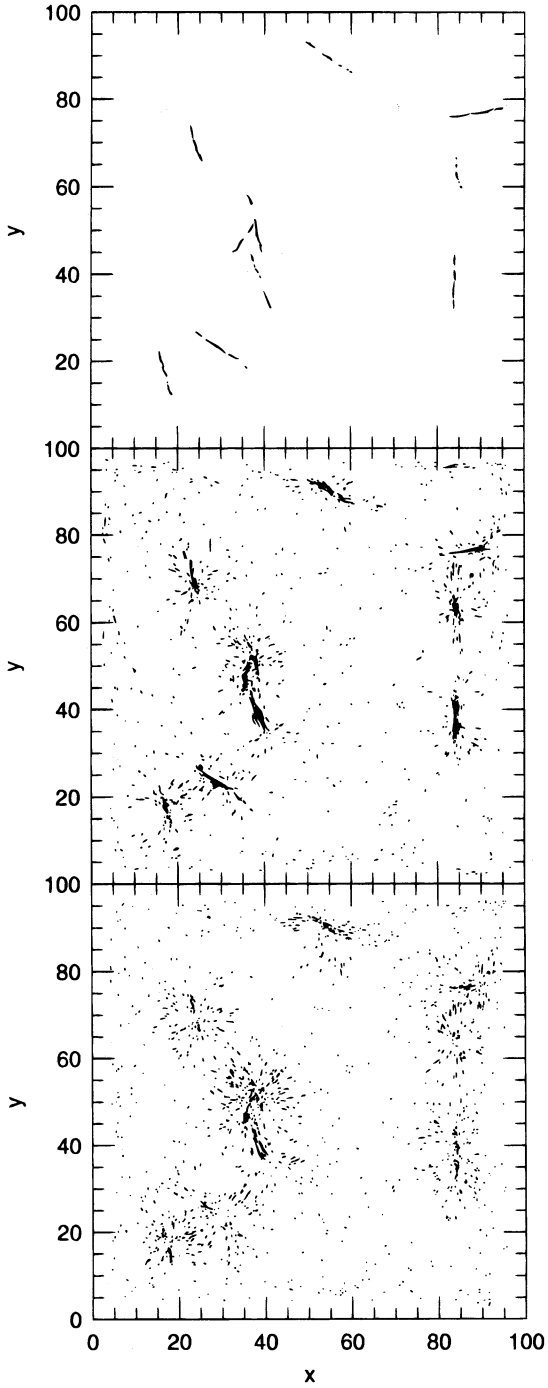

FIG. 2.—Directions of spindles with a window diameter of 5 (*top panel*), 20 (*middle panel*), and 25 (*bottom panel*). The lengths of the line segments are proportional to the signal I_I.

show strong signals with coherent orientation close to the orientation of the built-in spindles, while background noise shows weak signals with random orientation; the structures are best revealed when the window diameter is close to the thickness of the spindles.

These tests demonstrate that the combination of shape statistics and directions of eigenvectors of the moment-of-inertia tensor not only can detect clusters but also can help distinguish filaments and pancakes and reveal their orientation, as long as the window size is close to the thickness of the structures. Although we do not know beforehand the intrinsic scale of structures in a galaxy distribution, we can vary the window size to produce more than one map from the distribution. The map of structure orientation serves as a nice geometric supplement to the algebraic shape statistics.

3. APPLYING THE SHAPE STATISTICS TO OBSERVATIONS

3.1. *Data*

The observational data sets we will use here are the first two slices of the CfA survey provided by J. Huchra and the PPS survey provided by R. Giovanelli. The CfA data contain 1795 galaxies with velocities less than 15,000 km s^{-1}, in a 135° (right ascension) × 12° (declination) strip centered at 12^h30^m and 32°.5. The PPS survey covers a much larger region of the sky, a 120° (right ascension) × 50° (declination) strip centered at 1^h (right ascension) and 25° (declination). There are 6612 galaxies with velocities less than 15,000 km s^{-1} in this survey. Although the total number of galaxies in these two surveys is quite different, the number densities are about the same. Hence, their shape statistics can be compared directly. Figures 3 and 4 (*top panels*) show the distributions of galaxies in the two samples. The "great wall" is the most predominant feature in the CfA survey, while in the PPS survey, the coherent structures are of much smaller scales and there are more fingers of god produced by clusters.

3.2. *Shape Statistics of the Surveys*

The range of declinations for the CfA survey is only 12°, corresponding to about 20 h^{-1} Mpc at a distance of 100 h^{-1} Mpc. To study structures on scales larger than 20 h^{-1} Mpc, windows of these sizes are not completely embedded in the surveyed volume. To reduce edge effects, calibrating random samples are generated according to the geometry of the intersection of the window and the surveyed volume whenever a window crosses an edge of the volume.

Since both surveys are magnitude limited, the calibrating random samples should be generated according to a selection function determined from the observations. However, numerical tests of the shape statistics show that the difference between using a pure Poisson sample and using a Poisson sample modulated by a selection function is negligibly small. This is due to two facts. One is that the shape statistics, by construction, are insensitive to an overall constant gradient in a distribution. The other is that the Schechter form of the luminosity function is fairly smooth. To reduce computing time, we use Poisson samples to calibrate the signal. Numerical results show that 10 Poisson samples for each window are enough to reduce the variance of the shape statistics from random samples to a negligibly small level compared to the variance of the shape statistics of the data sample among different windows. As in Paper I, we use $\langle S_r \rangle$ to denote the mean value of the shape statistics over 10 calibrating samples. The signal from a window will be $S - \langle S_r \rangle$, where S is the shape statistics for the galaxies in a window from the observation. We then take the average of $S - \langle S_r \rangle$ from all windows to give the final signal of the shape statistics, $\langle S - \langle S_r \rangle \rangle$.

In Paper I, we described a method for estimating the variance of the mean shape statistics over windows in a sample of point distributions. We refer to this as the internal variance estimate since it is based on windows from one sample. We also ran nine simulations for a CDM model with $\Omega_0 = 1$ with different initial conditions and estimated

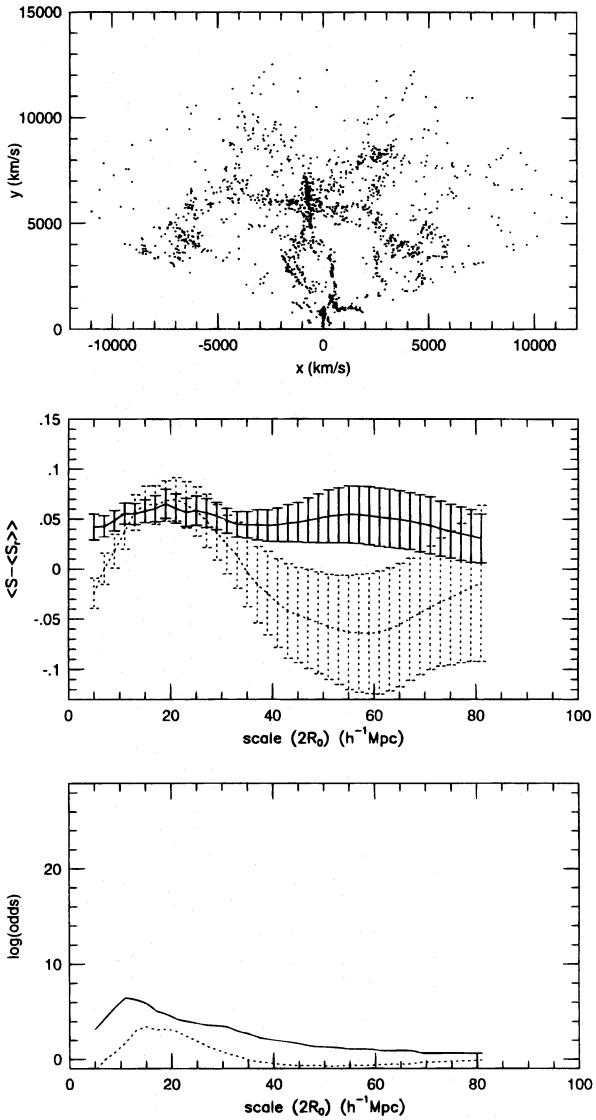

FIG. 3.—The CfA survey. (*top*) The distribution of 1795 galaxies in the CfA survey; (*middle*) the line (*solid line*) and plane (*dotted line*) shape statistics on different scales; (*bottom*) the logarithm (base 10) of the odds on a positive line (*solid line*) and plane (*dotted line*) shape statistics on different scales.

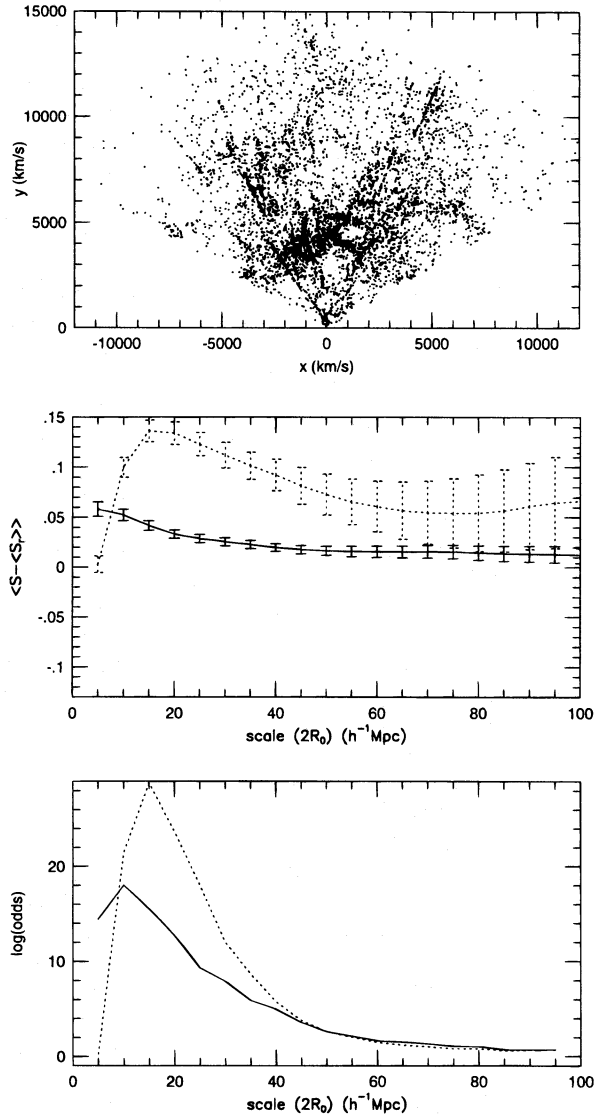

FIG. 4.—The PPS survey. (*top*) The distribution of 6612 galaxies in the PPS survey; (*middle*) the line (*solid line*) and plane (*dotted line*) shape statistics on different scales; (*bottom*) the logarithm (base 10) of the odds on a positive line (*solid line*) and plane (*dotted line*) shape statistics on different scales.

the variance of the shape statistics among different simulations. This is referred to as the external variance estimate. Numerical results show that the difference between these two estimates is within a factor of 2, with the internal estimate being greater than the external estimate, in good agreement with what was expected from the uncertainty of estimating the overlapping of neighboring windows, as described in Paper I. We adopt the internal estimate of the variance in this paper. Therefore, our estimation of the significance for a given signal is conservative.

Figure 3 (*middle panel*) shows the mean line (*solid line*) and plane (*dotted line*) shape statistics calibrated by Poisson samples for the CfA survey with different window diameters ($2R_0$). Error bars show 1 σ. Figure 3 (*bottom panel*) shows the logarithm (base 10) of the odds on a positive signal (i.e., odds on $\langle S - \langle S_r \rangle \rangle > 0$).

The mean value of the line shape statistic (in excess of that for a Poisson sample) of the CfA survey is more or less a constant all the way up to about 60 h^{-1} Mpc, with a broad peak around 20 h^{-1} Mpc. It then decreases gradually toward scales greater than 60 h^{-1} Mpc. The variance has a minimum around 10 h^{-1} Mpc. On smaller scales, the variance is larger due to the intrinsic variation of the distribution from window to window. It increases toward larger scales because of the smaller number of windows being used. Thus, the significance decreases toward large scales, with a peak around 10 h^{-1} Mpc. This scale corresponds to the thickness of the fingers of god in the CfA survey. As we will see shortly, on scales less than 10 h^{-1} Mpc, the finger-of-god effects contribute significantly to the signal; on scales much larger than 10 h^{-1} Mpc, the signal comes mainly from the "great wall."

The strong negative signal in the plane shape statistics on scales greater than about 35 h^{-1} Mpc is due to the small scale of the volume of the CfA survey in the declination direction, which is less than 20 h^{-1} Mpc. In a window with a large diameter (35 h^{-1} Mpc or larger) and thus intersecting with the surveyed volume, a uniform random distribution of points in such a planelike volume gives larger value for the plane shape statistic than a distribution with strong

clustering, and thus we have a negative value of $\langle S_p - \langle S_{pr} \rangle \rangle$. Even on smaller scales, the edge effects of such geometry greatly reduce the signal by producing large values of the plane shape statistic for the calibrating random samples. Therefore, because of the limited survey volume, the plane shape statistic is not appropriate for measuring structures in the CfA survey. This problem disappears in the PPS survey, which has large extension in both directions of right ascension and declination. On the other hand, the line statistic is not sensitive to a plane geometry, and the shape of the surveyed volume of the CfA survey has little effect on it.

Figure 4 (*middle panel*) and (*bottom panel*) shows the shape statistics and the logarithm of the odds on positive signals for the PPS survey. Because the declination range of the PPS survey is much larger than that of the CfA survey, the plane shape statistics show a real signal. There are many more galaxies in the PPS survey than in the CfA survey. Therefore, the variances of the shape statistics are much smaller, yielding much higher significance level for the signals. The plane shape statistic peaks at $\sim 15\ h^{-1}$ Mpc. On that scale, the filaments generated by the finger-of-god effect are well aligned to define a plane, resulting in a large value for the plane shape statistic. On smaller scales, a window is too small to cover two separated clusters, and the plane shape statistic goes down. On larger scales, the signal-to-noise ratio drops due to more field galaxies covered by a window. The line shape statistic of the PPS survey decreases monotonically as scale increases, indicating that the filamentary clustering is strongest on scales $\leq 5\ h^{-1}$ Mpc. On scales less than $10\ h^{-1}$ Mpc, the line shape statistic for the PPS survey has greater values than that for the CfA survey because there are more small scale clusters in the PPS survey which produce more fingers of god. On scales greater than $10\ h^{-1}$ Mpc, the line shape statistic for the PPS survey has smaller values than that for the CfA survey, due to the fact that there is no large-scale coherent structure in the PPS survey like the "great wall" in the CfA survey. Overall, the two surveys show significantly different structures on all scales considered here. It implies that neither of these two surveys is a fair presentation of the large-scale structure in the whole universe.

These results show that the shape statistics are efficient in distinguishing the difference between galaxy distributions on different scales.

3.3. *Maps of Structure Orientation*

To demonstrate the effectiveness of the shape statistics for detecting structures in a galaxy survey, we plot the map of structure orientation introduced in § 2 for the CfA survey and the PPS survey. Since the plane shape statistic is not a good measurement of pancakes for the CfA survey, we only plot the maps from the line shape statistic for the CfA survey in the top, middle, and bottom panels in Figure 5, with the diameters of windows being $10\ h^{-1}$ Mpc, $20\ h^{-1}$ Mpc, and $30\ h^{-1}$ Mpc, respectively. Figures 6 and 7 show the maps from the line and plane shape statistics for the PPS survey, using the same values of window diameter. The lengths of the line segments are proportional to the signal $I = (S - \langle S_r \rangle)/\sigma_r$, but the absolute normalizations from figure to figure are different.

From the top panel in Figure 5 we can see that the contribution to the line shape statistics in the CfA survey on scale of $10\ h^{-1}$ Mpc comes mainly from the finger-of-god effects

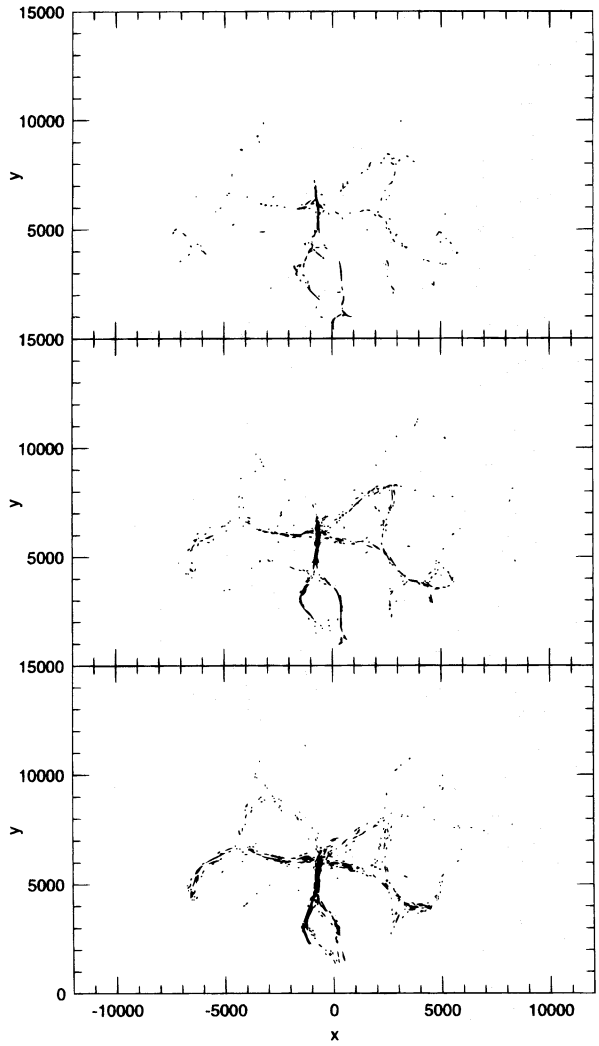

FIG. 5.—The map of structure orientation from the line shape statistic for the CfA survey with a window diameter of $10\ h^{-1}$ Mpc (*top panel*), $20\ h^{-1}$ Mpc (*middle panel*), and $30\ h^{-1}$ Mpc (*bottom panel*).

of the clusters, especially the Coma cluster. On larger scales, the contribution of the "great wall" becomes more and more important. In Figure 6, the PPS survey, on one hand, shows a lot more fingers of god but, on the other hand, has a much more complicated interconnected filamentary network compared to the CfA survey. This explains why the line shape statistics for the PPS survey have larger values than the CfA survey on small scales. With larger window sizes, the linear ridge discussed by Wegner, Haynes, & Giovanelli (1993) is clearly revealed, but the PPS survey still shows a rather complicated structure with shorter coherence scales than that in the CfA survey. This accounts for the fact that the line shape statistics have smaller values for the PPS survey than for the CfA survey on large scales. The orientation maps help visualizing the difference between the CfA survey and the PPS survey measured by the shape statistics on different scales.

In the maps for the PPS survey, the directions of line segments from the line shape statistic within a cluster are well defined, while those from the plane shape statistic are rather poorly defined. This indicates that the PPS survey is dominated by filamentary structure. As our toy model tests in Paper I have shown, a strong signal of plane shape statistic alone does not mean a strong signal for pancakes due

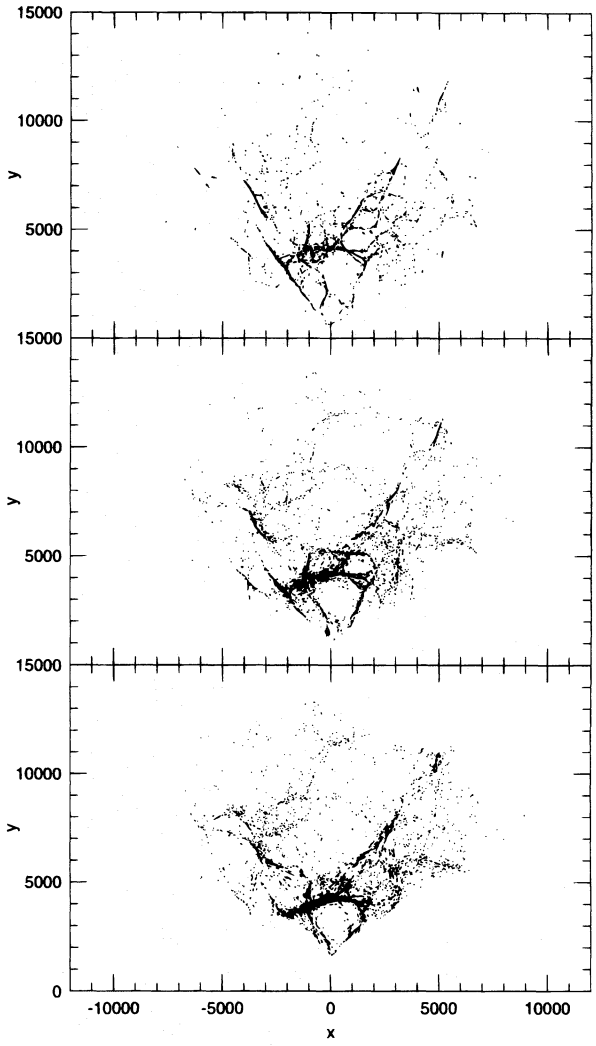

FIG. 6.—The map of structure orientation from the line shape statistic for the PPS survey with a window diameter of 10 h^{-1} Mpc (*top panel*), 20 h^{-1} Mpc (*middle panel*), and 30 h^{-1} Mpc (*bottom panel*).

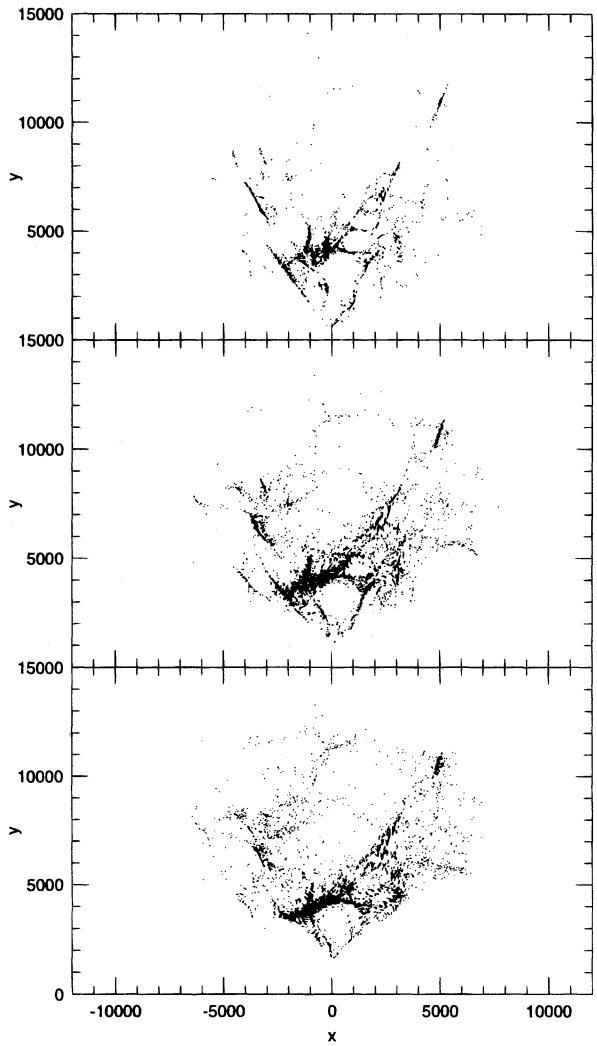

FIG. 7.—The map of structure orientation from the plane shape statistic for the PPS survey with a window diameter of 10 h^{-1} Mpc (*top panel*), 20 h^{-1} Mpc (*middle panel*), and 30 h^{-1} Mpc (*bottom panel*).

to the sensitivity of the plane shape statistic to poorly defined spindles. If the planelike structures suggested by a large value of the shape statistics were real, the orientation of the line segments in the cluster would be coherent. Therefore, the map of structure orientation helps distinguish true detections from false detections of planelike structures.

These maps show that the shape statistics can detect clusters in a galaxy survey on different scales. They also tell the nature of the structures, namely, whether they are filamentary or pancake-like. They reveal the contributions to the shape statistics from components on different scales in a distribution.

4. APPLYING THE SHAPE STATISTICS TO SIMULATIONS

4.1. *Simulation*

All N-body simulations presented in this paper were done using the Particle-Particle/Particle-Mesh (or P^3M) algorithm (Hockney & Eastwood 1981; Efstathiou & Eastwood 1981; Klypin & Shandarin 1983; Efstathiou et al. 1985, hereafter EDFW). The calculations evolve a system of gravitationally interacting particles in a cubic volume with triply periodic boundary conditions, comoving with Hubble flow. The forces on particles are computed by solving the Poisson equation on a $64 \times 64 \times 64$ grid using a fast Fourier transform method. The resulting force field represents the Newtonian interaction between particles down to a separation of a few mesh spacings. At shorter distances the computed force is significantly smaller than the physical force. To increase the dynamical range of the code, the force at short distance is corrected by direct summation over pairs of particles separated by less than some cutoff distance r_e. With the addition of this so-called *short-range correction*, the code accurately reproduces the Newtonian interaction down to the softening length η. In all calculations, η and r_e were set equal to 0.3 and 2.7 mesh spacings, respectively. With these particular values, the code has a dynamical range of 3 orders of magnitude (EDFW). The particular version of P^3M we used in this paper uses the so-called tilde coordinates (Shandarin 1980). The system is evolved forward in time using a Runge-Kutta time-integration scheme with a variable time step. We define a system of units by setting the mass $M_{\rm sys}$ of the system, the comoving side $L_{\rm box}$ of the computational volume, and the gravitational constant G equal to unity.

We consider two different cosmological models, the cold dark matter (CDM) model and the hot dark matter (HDM) model. For both models, we ran simulations with $\Omega_0 = 1$

and $\Omega_0 = 0.5$. We set the present value H_0 of the Hubble constant equal to 50 km s^{-1} Mpc^{-1}. In all cases, we set the comoving length of the computational volume $L_{box} = 200$ Mpc (present length units). The total mass of the system is $M_{sys} = 3H_0^2 \Omega_0 L_{box}^3/8\pi G = 5.55 \times 10^{17} \Omega_0\ M_\odot$. We use $32^3 = 32{,}768$ equal mass particles. The mass per particle is therefore $M_{part} = M_{sys}/32^3 = 1.694 \times 10^{13}\ M_\odot$ for the flat models and $3.388 \times 10^{12}\ M_\odot$ for the open models. These masses are comparable to the mass of a few galaxies. We therefore equate our particles with galaxies in what follows, although a more realistic approach would have to take into account the difference between the galaxy distribution and the mass distribution.

The method we use to set up initial conditions is fairly standard. We lay down $32^3 = 32{,}768$ particles on a uniform grid and displace them from their initial position in order to represent the initial density fluctuations. We then compute the initial peculiar velocities using the linear perturbation solution for a pure growing mode.

We assume that the initial fluctuations originate from a Gaussian random process. The initial density contrast can then be expressed as a superposition of plane waves with random phases:

$$\delta(\boldsymbol{x}) = \int \delta_{\boldsymbol{k}}^{\rm cont} e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} d^3k\ , \tag{3}$$

where $\delta_{\boldsymbol{k}}$ is the amplitude of the $\boldsymbol{k}$-mode. The integration in equation (3) is performed over all $\boldsymbol{k}$-space (hence the superscript "cont" that stands for "continuous"). Our simulations assume periodic boundary conditions. This restricts the range of possible values for the wavenumber $\boldsymbol{k}$ to multiples of the fundamental wavenumber $k_0 \equiv 2\pi/L_{box}$. We must therefore replace the integral in equation (3) by a sum over all modes that are allowed,

$$\delta(\boldsymbol{x}) = \sum_{\boldsymbol{k}} \delta_{\boldsymbol{k}}^{\rm disc} e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\ , \tag{4}$$

where the sum is over all values of $\boldsymbol{k} = (l, m, n)k_0$, with l, m, and n integers. The discrete and continuous amplitudes are related by

$$\delta_{\boldsymbol{k}}^{\rm cont} = \frac{V_{box}}{(2\pi)^3}\, \delta_{\boldsymbol{k}}^{\rm disc}\ , \tag{5}$$

where $V_{box}/(2\pi)^3 = (L_{box}/2\pi)^3 = k_0^{-3}$ is inverse of the density of discrete modes in $\boldsymbol{k}$-space. Notice that the discrete amplitude is dimensionless, while the continuous one has dimensions of (length)^3. The requirement that $\delta(\boldsymbol{x})$ is real implies $\delta_{\boldsymbol{k}} = \delta^*_{-\boldsymbol{k}}$.

The discrete amplitude is related to the power spectrum by

$$P(k) = \frac{V_{box}}{(2\pi)^3} |\, \delta_{\boldsymbol{k}}^{\rm disc}\,|^2\ . \tag{6}$$

The power spectrum is often expressed as

$$P(k) = AkT(k)^2\ , \tag{7}$$

where A is the amplitude and has dimension of (length)^4 and $T(k)$ is the transfer function. In most models, $T(k)$ goes to unity at high redshift, and we recover the Harrison-Zeldovich power spectrum $P(k) \propto k$. The value of the amplitude is fixed by the value of the cosmic microwave background temperature anisotropy, as measured by *COBE* (Smoot et al. 1992),

$$A = \frac{1}{(2\pi)^2}\frac{6\pi^2}{5} Q_2^2 R_H^4\ , \tag{8}$$

where Q_2 is the temperature quadrupole anisotropy and R_H is the radius of the horizon. For all simulations, we used the value $A = 6.54 \times 10^5\ h^{-4}$ Mpc$^4 = 1.0464 \times 10^7$ Mpc^4 given by Bunn, Scott, & White (1995) for standard CDM models (that value was published in a preprint; the actual value given in their final paper was $A = 8.16 \times 10^5\ h^{-4}$ Mpc^4).

For the CDM simulations, we used the transfer function given by Bardeen et al. (1986):

$$T(k) = \mathscr{L}(z)\frac{\ln(1 + 2.34q)}{2.34q}[1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4]^{1/4}\ , \tag{9}$$

where $\mathscr{L}(z)$ is the linear growth factor between the initial state and the present [$\mathscr{L}(z) = 1 + z$ for $\Omega = 1$ models], $q = k\theta^{1/2}/(\Omega_x\, h^2$ Mpc), with Ω_x being the density parameter of the dark matter (nonbaryonic) component, which we set to 15/16, and $\theta = 1$ for models with three flavors of relativistic neutrinos. For the HDM models, we used

$$T(k) = \mathscr{L}(z) \times 10^{-(k/k_\nu)^{1.5}}\ , \tag{10}$$

where k is in units of Mpc^{-1} and $k_\nu = 2\pi/\lambda_\nu$, where $\lambda_\nu = 56$ Mpc is the damping length of the neutrinos (Kolb & Turner 1990).

Once these power spectra are specified, we can compute the initial positions of the particles. We lay down particles on a uniform grid $32 \times 32 \times 32$ and then displace each particle according to

$$\Delta\boldsymbol{x} = -\sum_{\boldsymbol{k}} \frac{|\,\delta_{\boldsymbol{k}}^{\rm disc}\,|\,\boldsymbol{k}}{2\pi k^2} \sin(2\pi\boldsymbol{k} - \phi_{\boldsymbol{k}})\ , \tag{11}$$

where $\phi_{\boldsymbol{k}}$ is a random number between 0 and 2π. The initial peculiar velocities are given by

$$\boldsymbol{v}_i = \frac{\dot{b}(t)}{b(t)}\Delta\boldsymbol{x}\ , \tag{12}$$

where $b(t)$ is the linear growing mode of the perturbation.

Figures 8*a*–8*d* show the galaxy distributions at redshift of zero for the four models, namely, C10 (CDM model with $\Omega_0 = 1$), H10 (HDM model with $\Omega_0 = 1$), C05 (CDM model with $\Omega_0 = 0.5$), and H05 (HDM model with $\Omega_0 = 0.5$). We used the same ensemble of random phases $\phi_{\boldsymbol{k}}$ for all models.

The general visual impression of these simulations is that the CDM models have stronger clustering on small scales, while the HDM models have structures with greater coherent length. The models with $\Omega_0 = 1$ are more clustered than the models with $\Omega_0 = 0.5$. However, the small clusters in the CDM models more or less outline the structures (presumably pancakes) of large sizes in the HDM models. A similar result was obtained by Melott & Shandarin (1993) in a series of N-body simulations. This feature is not obvious to the eye when we look at the galaxy distributions of the CDM models alone, but it becomes quite clear when we compare the galaxy distribution of a CDM model with its HDM counterpart side by side. We will see below that this is detected by the shape statistics, demonstrating that

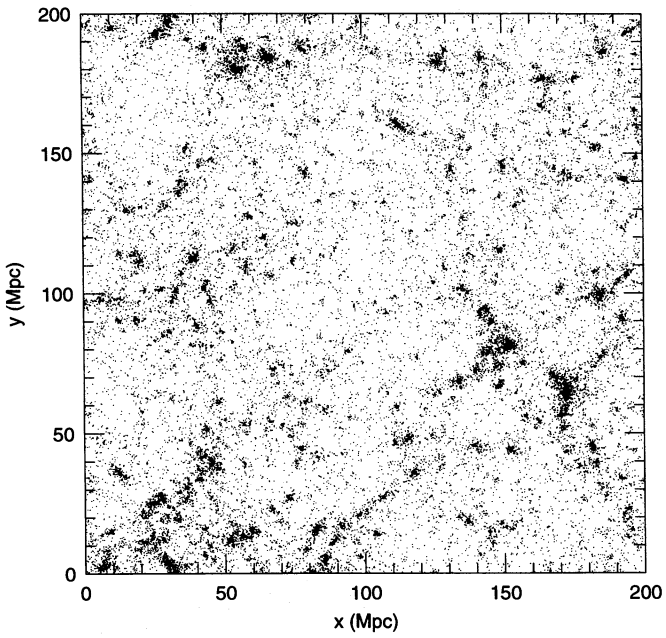

FIG. 8*a*

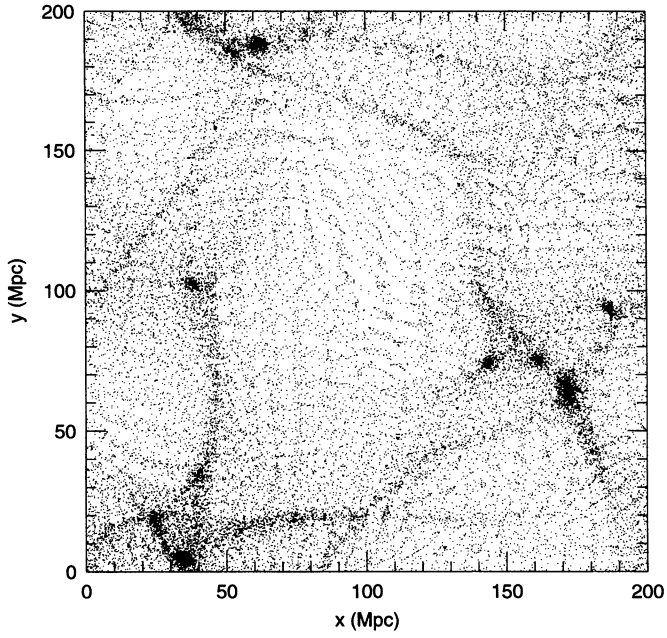

FIG. 8*b*

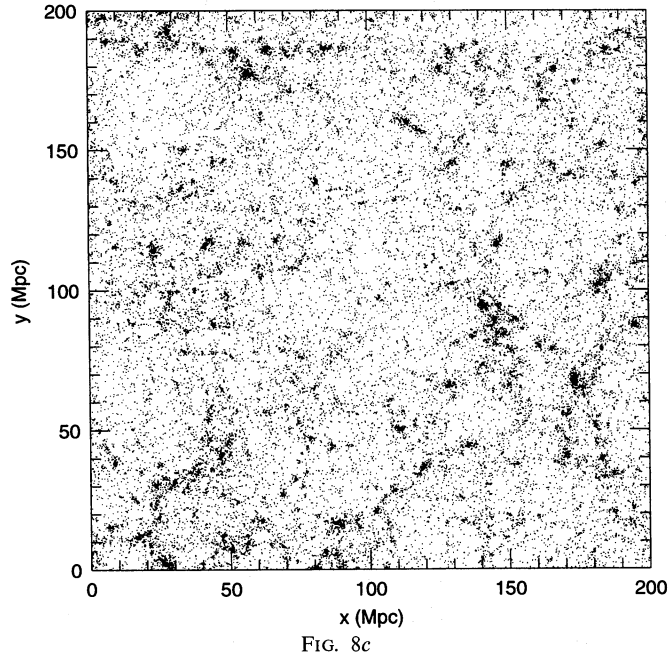

FIG. 8*c*

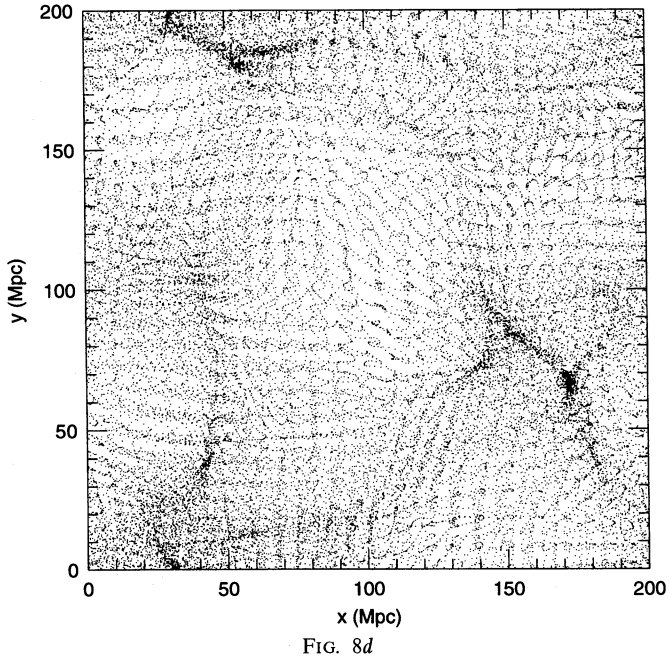

FIG. 8*d*

FIG. 8.—The distributions of 32^3 galaxies at redshift $z = 0$ from (*a*) a CDM model with $\Omega_0 = 1$; (*b*) an HDM model with $\Omega_0 = 1$; (*c*) a CDM model with $\Omega_0 = 0.5$; (*d*) an HDM model with $\Omega_0 = 0.5$.

the statistics can successfully detect features that could easily be missed by visual inspection.

It is also apparent that, even at redshift of zero, a large number of points in the HDM models are still close to their initial positions. The regularity of the initial condition (the lattice) is largely preserved. This indicates that the low-density regions are still in the linear regime of gravitational evolution. The shape statistics can "see" this regularity as well.

4.2. *Shape Statistics of Different Models*

To reduce computing time, we only use one-sixth of the galaxies in a simulation to compute the shape statistics. Compared with a full sampling of 32^3 galaxies, this sparse sampling tends to give a smaller value of $\langle S - \langle S_r \rangle \rangle$, since $\langle S_r \rangle$ is greater for a smaller number of galaxies in a window. Also the variances tend to be larger due to the smaller number of windows used. Therefore, the signal is weaker than that for a full sampling. Nevertheless, as we will see below, it is still strong enough for us to extract useful information about clustering in the simulations.

Figures 9*a*–9*b* show the line and plane shape statistics, respectively, at redshift of zero as a function of window diameter for the four models. Error bars show 1 σ. We notice that the shape statistics for $\Omega_0 = 1$ models (C10 and H10; solid and dotted lines, respectively) are generally

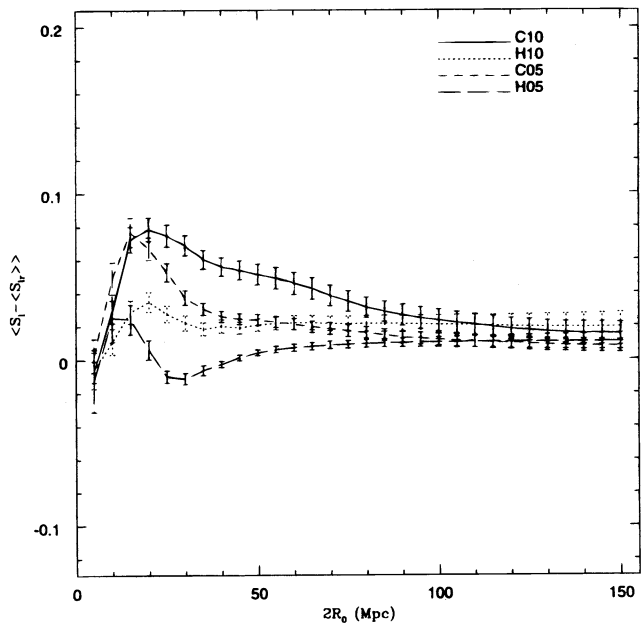

FIG. 9a

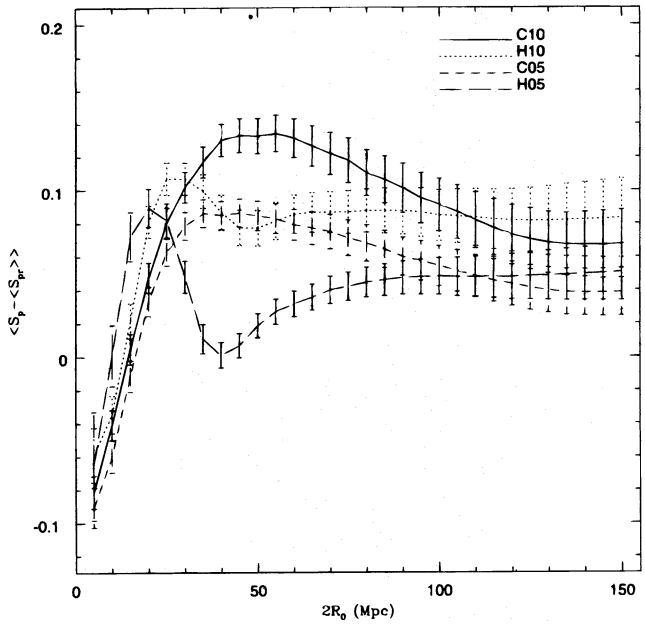

FIG. 9b

FIG. 9.—The shape statistics of galaxy distributions from four models at redshift $z = 0$. (*a*) The line shape statistics; (*b*) the plane shape statistics.

larger than those for $\Omega_0 = 0.5$ counterparts (C05 and H05; short dashed and long dashed lines, respectively). This is in good agreement with the fact that the structures are much more evolved and the clustering is stronger in the $\Omega_0 = 1$ models due to stronger gravitational interaction.

Our tests in Paper I show that the shape statistics peak on scales of a few times the typical scale of the underlying structures. The cluster size in the $\Omega_0 = 1$ models is generally greater than that in the $\Omega_0 = 0.5$ models. Except for the plane shape statistics for the CDM models, the shape statistics for the $\Omega_0 = 0.5$ models generally peak on a smaller scale. They also drop faster than their $\Omega_0 = 1$ counterparts as the scale increases because galaxy distributions in the $\Omega_0 = 0.5$ models approach a homogeneous distribution on scales smaller than those in the $\Omega_0 = 1$ models. The plane shape statistics for the CDM models behave differently because the structures in a CDM model are clusters and filaments. There is no well-defined planelike structure in a CDM model. Therefore, there are no well-defined peaks in the curve for the plane shape statistics of the CDM models.

Another predominant feature in Figure 9 is that the shape statistics of the CDM models *decrease* from intermediate scales to the largest scales, while the shape statistics of the HDM models *increase* from intermediate scales to the largest scales. It is a clear signature of the difference between these two kinds of models. The coherence length in a HDM model is comparable to the size of the simulation box, while that in a CDM model is much smaller. With a CDM model, when the diameter of windows increases, more field galaxies and clusters of random orientation are included in a window, and thus the signal is gradually washed away by increasing background noise. On the other hand, the signal from a HDM model remains strong since a larger window also covers more of a pancake due to the large size of the pancakes, and thus the signal-to-noise ratio remains more or less the same. This demonstrates that the two families of models (CDM and HDM) have significantly different behavior in terms of the global dependence of the shape statistics on window diameter.

From Figure 9*b* we see that the plane shape statistics for the CDM models are greater than their HDM counterparts over a large range of scales. This is caused by the large-scale alignment of the small-scale filaments in the CDM models we mentioned in the previous section. As was demonstrated in Paper I, two aligned filaments being covered by one large window form a plane and produce a large signal for the plane shape statistics. Figures 4*b* and 5*b* in Paper I show larger plane shape statistics on large scales for model A4 consisting of spindles than for model A5 consisting of pancakes. On the other hand, the plane shape statistics for A5 are larger than A4 on scales a few times the thickness of the built-in pancakes. In our simulations, the combination of the alignment of filaments in the CDM models mimicking the structure of pancakes in the HDM models and the high density of the clusters being aligned produces large values for the shape statistics. However, on scales less than about 20-30 Mpc the plane shape statistics for the HDM models are greater than those for the CDM models, as shown in Figure 9*b*. On such scales, the window size is a few times of the thickness of pancakes in a HDM model, and this gives a strong signal for the plane shape statistic. On the other hand, the signal of the plane shape statistic from a group of clusters in a CDM model is small. Therefore, CDM models and HDM models have rather different behavior on all scales in terms of the shape statistics.

At first sight, it is counterintuitive to see a stronger plane shape statistic for a CDM model than for an HDM model. One would expect otherwise because there is more power on intermediate to large scales in an HDM model. Our tests demonstrate that the morphology of large-scale structure is not related to the power spectrum of density perturbations in as simple a way as the two-point correlation function, which is not surprising since only for a Gaussian field does the two-point correlation function (or equivalently the power spectrum) reveal all the information about the underlying field. The shape statistics involve up to the six-point correlation function and thus carry much more information about the large-scale structure.

Figure 9 also shows that the shape statistics are negative on very small scales. This is because on such scales the clusters are more or less spherical. They produce shape statistics less than those from a Poisson sample and yield a negative $\langle S - \langle S_r \rangle \rangle$. This effect is stronger for the plane shape statistics since $\langle S_{pr} \rangle > \langle S_{lr} \rangle$, while both $\langle S_p \rangle$ and $\langle S_l \rangle$ are small.

The dip in the shape statistics of the HDM models is due to the semiperiodicity of the field galaxies preserved from the initial condition, as we saw visually in Figures 8*b*–8*d*. A distribution of points "more uniform" than a Poisson distribution will produce a negative signal for the shape statistics since $\langle S \rangle < \langle S_r \rangle$. A window with diameter a few times the period of the regular distribution has the highest sensitivity to the regularity. On very small scales, a window cannot cover a unit of the periodic structure, and, on large scales, such regularity is averaged out.

Our results demonstrate that the dependences of the shape statistics on scale are very different for the CDM and HDM models. The shape statistics are also rather different for models with different values of Ω_0. They are very powerful tools to discriminate different models for structure formation.

It is interesting to compare the shape statistics from the simulations with those from the observations. We see from Figure 3*b* and Figure 9*a* that the shape of the curve for the line shape statistic for the CfA survey is rather different from those for models C10, C05, and H05, but quite similar to that for model H10. On the other hand, the shape statistics for the PPS survey are similar to those for CDM models as shown in Figure 4*b* and Figure 9. The fact that the CfA survey and the PPS survey seem to suggest quite different origins for the observed structures implies that they are not fair samples of the universe. Therefore, the various challenges to the standard model based on current redshift surveys should be taken with caution. Deeper redshift surveys are needed to provide a fair presentation of the universe. Only then could we pin down the correct model for structure formation.

4.3. *Evolution of Shape Statistics*

In the previous subsection, we saw that the shape statistics are sensitive to the difference between the morphologies of the structures produced by CDM and HDM models. They also depend on the value of Ω_0. Thus, they are powerful tools to discriminate theories of structure formation. In this subsection, we investigate the possibility of using the shape statistics to trace the dynamic evolution of structure formation. We compute the shape statistics for the simulations of the four models at different redshifts to study their time evolution. Our results show that the $\Omega_0 = 0.5$ models have similar behavior to that of the $\Omega_0 = 1$ models. We only present the results for the $\Omega_0 = 1$ models (i.e., models C10 and H10) here.

Figures 10*a*–10*b* show the mean line and plane statistics, $\langle S - \langle S_r \rangle \rangle$, of simulations of model C10 at redshifts of 0, 0.05, 0.1, 0.2, 0.5, 1, and 2 as a function of the diameter of the windows. Error bars are 1 σ. From these figures we see in the curve for $z = 2$ a dip in the shape statistics at $2R_0 \sim 30$ Mpc, a few times the mean distance between galaxies in the simulation. This negative signal is produced by the semiregularity in the galaxy distribution inherited from the initial condition. As clustering evolves, the regularity is erased gradually, and the dip disappears in the curve for $z = 1$. For distributions at high redshifts, the standard deviation of the shape statistics on small scales is rather large because at high redshifts clustering is very weak and, with a small size, only a few windows contain more than two points to yield meaningful shape statistics. Actually, on scales smaller than 20 Mpc, most of the windows contributing to the shape statistics contains only two or three points. These windows would give large values for the shape statistics if the line (plane) formed by the two (three) points happens to pass through the center of the window. There-

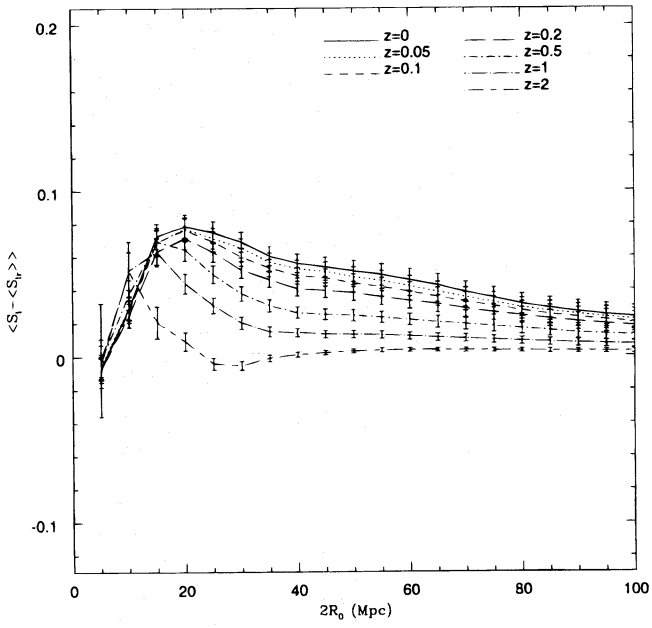

FIG. 10*a*

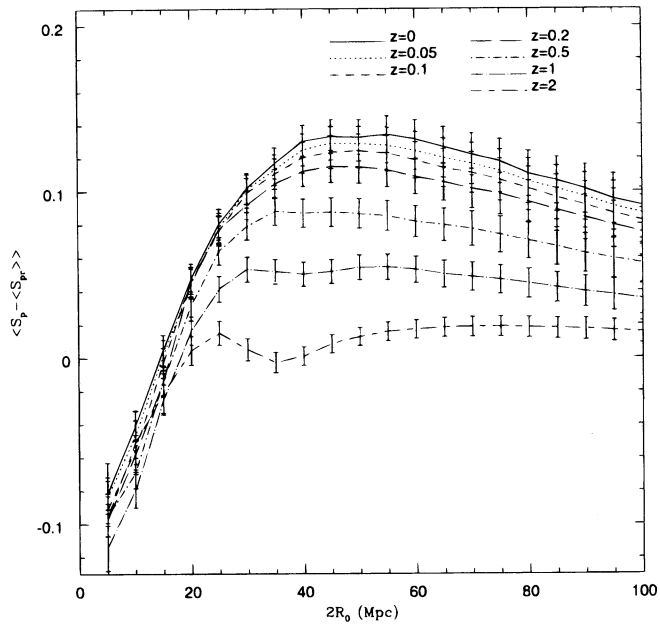

FIG. 10*b*

FIG. 10.—The shape statistics at different redshifts for the simulation of a CDM model with $\Omega_0 = 1$. Error bars show 1 σ. (*a*) The line shape statistics; (*b*) the plane shape statistics.

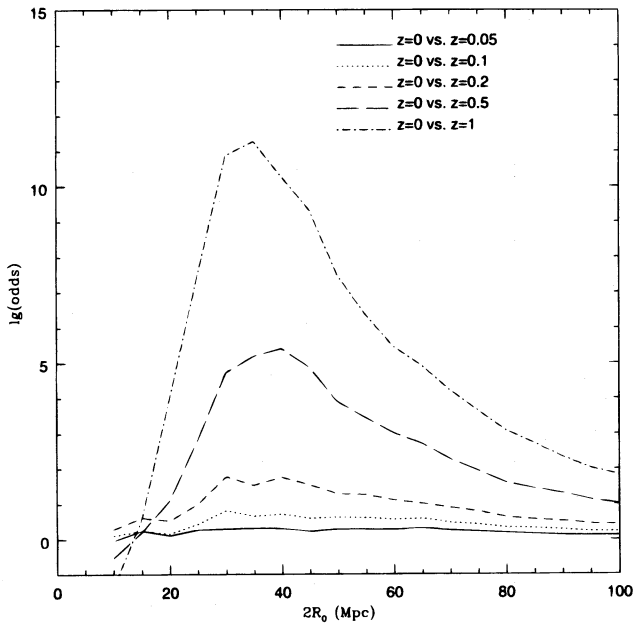

FIG. 11*a*

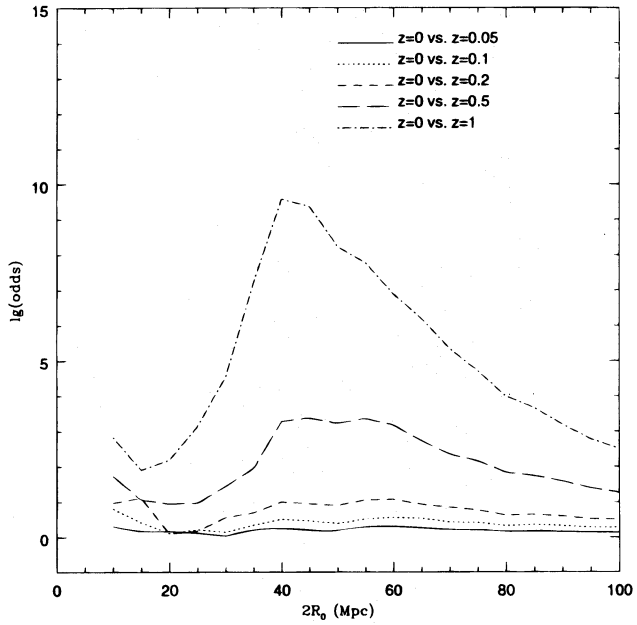

FIG. 11*b*

FIG. 11.—The logarithm (base 10) of the odds on the evolution of the shape statistics between redshifts of 0 and 0.05, 0 and 0.1, ..., and 0 and 1, for the simulation of a CDM model with $\Omega_0 = 1$. (*a*) The line shape statistics; (*b*) the plane shape statistics.

fore, on such small scales, the shape statistics of galaxy distributions at high redshifts tend to be the greatest with large uncertainty.

In Figure 10*a*, the peak in the curve moves toward large scales from $z = 2$ to $z = 0$, indicating that the size of clusters in the simulation grows as the structure evolves. The width of the peak also increases as the structure evolves from high redshifts to low redshifts, an indication that the distribution of the size of the clusters is broadened and more and more clusters of different sizes are generated during dynamical evolution.

During gravitational evolution, more structures form and the density contrast increases, and thus the shape statistics increase monotonically on scales greater than 20 Mpc as clustering evolves. Figures 11*a* and 11*b* show the significance of the evolution, that is, the logarithm of the odds that $\langle S - \langle S_r \rangle \rangle$ at $z = 0$ is greater than $\langle S - \langle S_r \rangle \rangle$ at $z = 0.05$, 0.1, 0.2, and so on. It is obvious from these figures that the shape statistics evolve significantly during structure formation. Even from the distribution at $z = 0.05$ to that at $z = 0$, the odds on an increased signal are greater than 1 on all scales greater than 15 Mpc. The morphological differ-

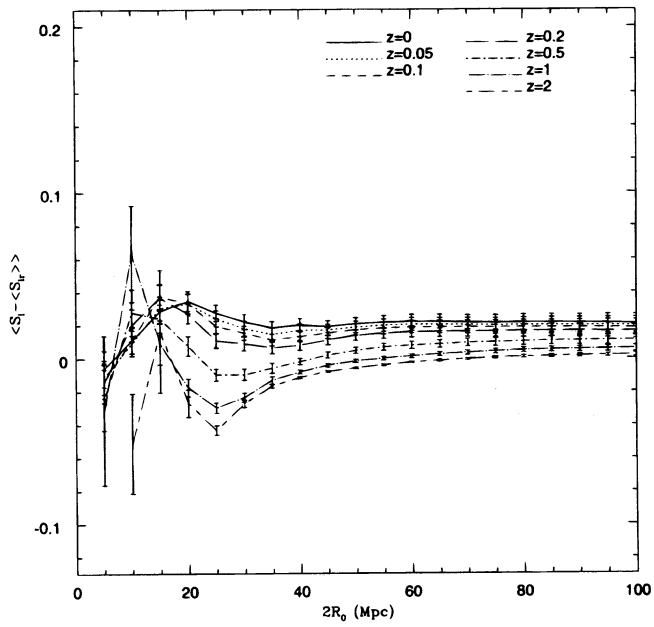

FIG. 12*a*

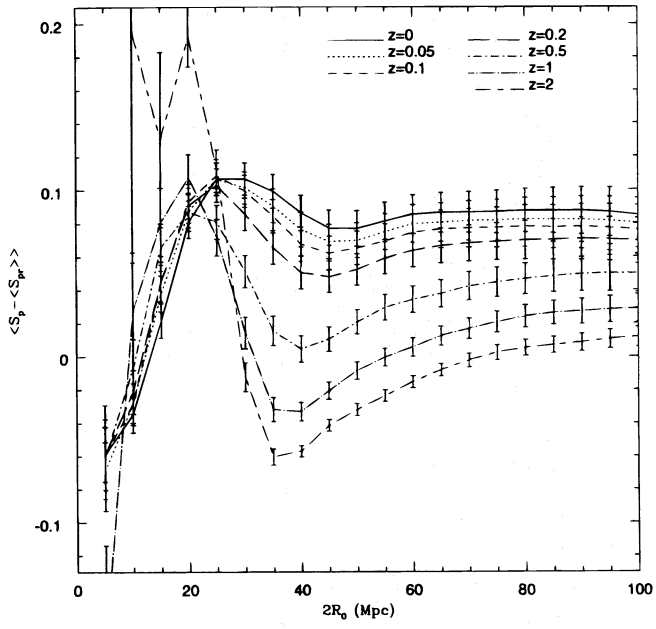

FIG. 12*b*

FIG. 12.—The shape statistics at different redshifts for the simulation of an HDM model with $\Omega_0 = 1$. Error bars show 1 σ. (*a*) The line shape statistics; (*b*) the plane shape statistics.

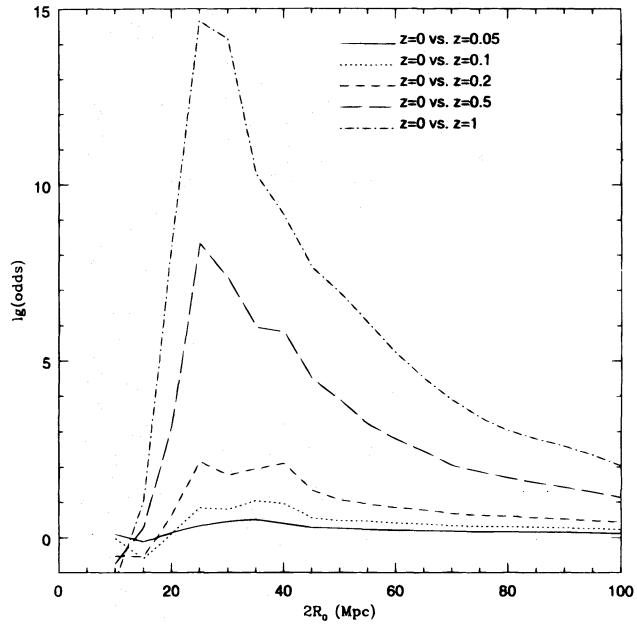

FIG. 13*a*

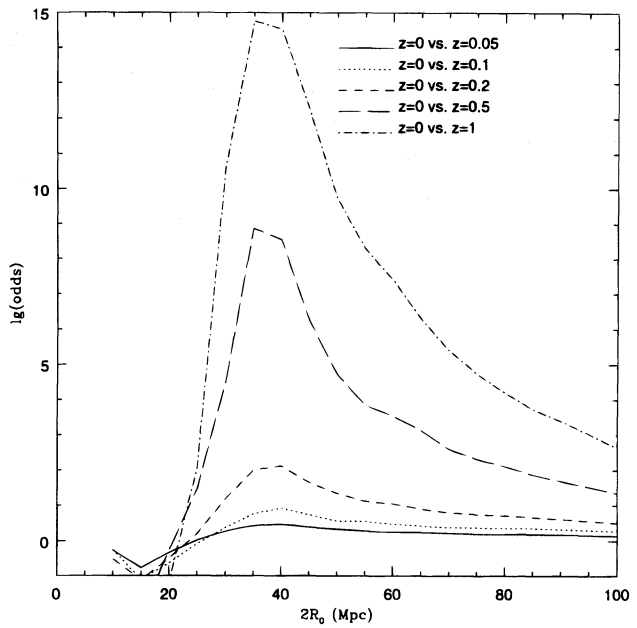

FIG. 13*b*

FIG. 13.—The logarithm (base 10) of the odds on the evolution of the shape statistics between redshifts of 0 and 0.05, 0 and 0.1, ..., and 0 and 1, for the simulation of an HDM model with $\Omega_0 = 1$. (*a*) The line shape statistics; (*b*) the plane shape statistics.

ence between galaxy distributions at $z = 0$ and $z = 0.2$ is readily detected by the shape statistics.

These results are in good agreement with the bottom-up picture presented by the CDM scenario. Gravity first generates clusters on small scales from some initial density fluctuation. Then the clusters grow, and the galaxy distribution becomes more and more filamentary toward larger scales in a hierarchical manner.

Figure 12 and Figure 13 are for the shape statistics for model H10. The general features of the evolution of the shape statistics are similar to those of model C10. The difference is that for model H10, the plane shape statistics peak on scale around 20 Mpc, while there is no such peak with model C10. This is due to the planelike structure in a HDM model. The peak moves to larger scales as the structure evolves, and the peak width is also broadened.

Our results show that the shape statistics are very sensitive to the morphological difference caused by clustering evolution of a galaxy distribution. They can thus be applied to surveys at different redshifts to study the evolution of gravitational dynamics. It is important to remember that these results may be affected by our failure to include biasing, especially at epochs when galaxy formation is still an ongoing process.

5. SUMMARY

In this paper, we have shown an improved way of presenting structures detected by the shape statistics in a galaxy distribution. The new method presents the orientation of detected structure units. It can reveal structures with less scatter than the old method in Paper I and helps distinguish filaments from pancakes. It gives a way to visualize the difference between galaxy distributions detected by the shape statistics and serves as a geometric supplement to the algebraic shape statistics. We have applied the shape statistics to the CfA survey and the PPS survey, as well as numerical simulations of CDM and HDM models with $\Omega_0 = 1$ and 0.5. Our results show that the shape statistics are very sensitive tools for detecting clusters in galaxy surveys on different scales. They are efficient in revealing differences between galaxy samples. We find that the structures in the CfA and PPS surveys are significantly different on all scales we consider. The PPS survey has stronger clustering on small scales, consistent with CDM models. The CfA survey has coherent structures on very large scales, consistent with HDM models. This suggests that neither of these two surveys is a fair sample of the universe. Therefore, challenges to the standard model for structure formation based on current redshift surveys should be taken with caution. Of course, the fact that such a large variance exists on the scale of current surveys may serve as a useful constraint on structure formation models. Redshift surveys on scales much larger than $100\ h^{-1}$ Mpc are needed to resolve the problem of structure formation. Our results also demonstrate that shape statistics are sensitive to the difference between the morphology of structures from CDM and HDM models with different density parameters. They trace the clustering evolution faithfully from a few Mpc to over 100 Mpc. Therefore, the shape statistics are very useful to constrain theories of structure formation. They can also serve as tools to study gravitational evolution of the large-scale structure. On the other hand, from the point of view of observational cosmology, they can detect and measure the amount of filaments and pancakes in galaxy surveys systematically and quantitatively.

We want to thank John Huchra for providing the data of the first two slices of the CfA survey, and Riccardo Giovanelli for providing us the data of the Pisces-Perseus Supercluster survey. The simulations were done at High Performance Computing Facility at University of Texas at Austin. This work is supported in part by NSF grants AST 90-20757, AST 91-19475, and AST 93-15375, NASA grants NAGW-2399 and NAG 5-2785, and Robert A. Welch grant F-1115.

REFERENCES

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Bunn, E. F., Scott, D., & White, M. 1995, ApJ, 441, L9
Efstathiou, G., Davis, M., Frenk, C. S., & White, S. D. M. 1985, ApJS, 57, 241 (EDFW)
Efstathiou, G., & Eastwood, J. W. 1981, MNRAS, 194, 503
Hockney, R. W., & Eastwood, J. W. 1981, Computer Simulation Using Particles (New York: McGraw-Hill)
Klypin, A., & Shandarin, S. F. 1983, MNRAS, 204, 891
Kolb, E. W., & Turner, M. S. 1990, The Early Universe (Redwood City: Addison-Wesley)
Luo, S., & Vishniac, E. T. 1995, ApJS, 96, 429
Melott, A. L., & Shandarin, S. F. 1993, ApJ, 410, 469
Shandarin, S. F. 1980, Astrophysika, 16, 769
Smoot, G. F., et al. 1992, ApJ, 396, L1
Wegner, G., Haynes, M. P., & Giovanelli, R. 1993, AJ, 105, 1251