THE ASTROPHYSICAL JOURNAL, 413:48–58, 1993 August 10

PERCOLATION TECHNIQUE FOR GALAXY CLUSTERING

ANATOLY KLYPIN AND SERGEI F. SHANDARIN
Department of Physics and Astronomy, University of Kansas, Lawrence, KS 66045
Received 1992 December 11; accepted 1993 February 18

ABSTRACT

We study percolation in mass and galaxy distributions obtained in three-dimensional simulations of the CDM, C + HDM, and the power law ($n = -1$) models in the $\Omega = 1$ universe. Percolation statistics is used here as a quantitative measure of the degree to which a mass or galaxy distribution is of a filamentary or cellular type. We have developed a very fast code (based on the 1985 algorithm described by Stauffer) which calculates the statistics of clusters along with the direct detection of percolation. We found that two parameters—μ_∞ (eq. [3]), characterizing the size of the largest cluster, and μ^2 (eq. [4]), characterizing the weighted mean size of all clusters excluding the largest one—are extremely useful for evaluating the percolation threshold. An advantage of using these parameters is their low sensitivity to boundary effects. We show that both the CDM and the C + HDM models are extremely filamentary both in mass and galaxy distribution. The percolation thresholds for the mass distributions are $p_c = 0.023 \pm 0.005$ in the C + HDM and $p_c = 0.044 \pm 0.005$ in CDM models compared to $p_c = 0.16$ for a Gaussian random field. For galaxy samples with a few thousand galaxies the thresholds are $p_{c,\mathrm{C+HDM}} = 0.06 \pm 0.02$ and $p_{c,\mathrm{CDM}} = 0.10 \pm 0.02$ compared to $p_c = 0.31$ for a Poisson distribution. Percolation in regions having the shape of a parallelepiped is discussed in the context of the applications of percolation statistics to real galaxy catalogs.

Subject headings: cosmology: theory — dark matter — galaxies: clustering — large-scale structure of universe — methods: numerical

1. INTRODUCTION

Angular and spatial distributions of galaxies show structures which are often referred to as cellular or filamentary. The former reflects a visual impression that galaxies are concentrated to narrow walls separating large isolated voids of galaxies and the latter means that galaxies are concentrated to one-dimensional threads forming a kind of three-dimensional web. No statistics which can quantitatively and unambiguously measure these impressions has been suggested despite many attempts.

Zel'dovich (1982) was the first to realize that this is a topological question. He suggested the following explanation of the formation of the cellular structure in the pancake scenario associated (at that time) with the neutrino-dominated (30 eV) universe. He argued that if a few percent of the volume is filled randomly with some substance and the rest of the volume remains empty, then the filled regions will be isolated and the empty space will look like a single ocean. (Of course, lakes on islands are also possible.) In order to make a few percent of the filled volume look like a connected structure one must provide a special arrangement of the filled regions in space. To support this line of reasoning quantitatively he suggested using percolation statistics. Soon after, one of the authors of this paper suggested using the percolation statistics as a descriptor of the observational distribution of galaxies and as a cosmological test (Shandarin 1983). The later development of this method was recently reviewed by Dominik & Shandarin (1992), and we will not repeat it here.

Instead we briefly describe the main idea of the percolation technique and its relation to the topology of the structure. In this paper we deal with three-dimensional cubic lattices in parallelepiped-like regions and define percolations in such systems. For example, let us take a three-dimensional cubic lattice of size N^3 and assign the labels "filled" to some of the cells and "empty" to all the rest according to some specified rule. For instance, if the density is given on the lattice, one can label the cells with the density higher than a specified threshold as filled and the others as empty. After assigning the labels each cell becomes a member of a cluster. Each cluster consists of the cells of one type which satisfy the requirement of the neighborhood. Two cells of the same type are considered to be neighbors if (1) they have a common side (at most six cells can be the neighbors of a cell due to this requirement) or (2) satisfy the principle: the neighbor of my neighbor is my neighbor. Sometimes the first requirement is modified, and the cells having a common ridge or corner are also included in the list of the immediate neighbors (Mo & Börner 1990; de Lapparent, Geller, & Huchra 1991). These modifications, without changing the principal idea of percolation analysis, do change the values of the percolation thresholds. One disadvantage of such modifications is that they have not been studied theoretically. In particular the percolation thresholds in Poisson and Gaussian distributions are not known.

The numbers of cells in a cluster is called the size or volume of the cluster and, generally speaking, can be any number from 1 to N^3. Percolation theory studies the transition of an infinite system ($N \to \infty$) from the state in which every cluster is finite, to one in which an infinite cluster exists. In finite systems, a cluster which spans the entire region plays the role of the infinite cluster. For example, in a finite cubic region such a cluster connects the antipodal sides. The formation of such a cluster may be considered a phase transition (e.g., from insulator to conductor).

The phenomenon of percolation can be observed in a variety of systems: various lattices (simple cubic, body centered cubic, face centered cubic, etc.), various dimensions, discrete or con-

tinuous systems. However, the essence of percolation is a sudden formation of the infinite cluster as some parameter characterizing the system changes gradually.

The percolation technique applied to galaxy distribution is very different from that used in solid state physics (see, e.g., Ziman 1979) or, say, for models of solar activity (Wentzel & Seiden 1992). In both cases the percolation phenomena models some physical process. The application of percolation analysis to galaxy distributions suggests using percolation statistics as a descriptor, although the connection with the physical processes of gravitational instability can also be traced (Zel'dovich 1982; Shandarin & Zel'dovich 1989). Two aspects of this application are worth emphasizing. First, one may expect that percolation statistics can be an objective quantitative measure of filamentary and/or cellar structures, and second, the percolation statistics can be an additional objective discriminator between various cosmological scenarios of the formation of the large-scale structure in the universe. These two aspects are relatively independent because the former emphasizes the difference with "structureless" distributions like Poisson and Gaussian and the latter the difference between the distributions in question (see the discussion below). In this paper we address both issues.

Percolation thresholds are primarily topological characteristics of random fields and are related to the Euler characteristic. In the case of Gaussian random fields in one-, two-, and three-dimensional spaces the genus curve changes the sign at percolation limits (Tomita & Murakami 1988) which agrees with Ziman's (1979) speculation.

The application of percolation analysis to real galaxy catalogs encounters several problems mentioned in the early studies (Shandarin 1983; Bhavsar & Barrow 1983; Einasto et al. 1984; Dekel & West 1985; Klypin 1987; Mo & Börner 1990; de Lapparent et al. 1991; Dominik & Shandarin 1992). Discreteness, typically complex shapes of samples, inhomogeneity related to the selection function, and small statistics are among them. In this paper we discuss the problems of discreteness and boundaries of the samples using simulated three-dimensional samples of galaxies in CDM, C + HDM, and the $n = -1$ power law models. We suggest a new technique to deal with the discreteness and also test three methods of estimating the percolation thresholds, two of which are quite insensitive to the boundary effects. We reserve the study of the inhomogeneous samples to a separate paper.

Concluding the introductory remarks we would like to emphasize a few general points:

1. Percolation is a phase transition characterizing a system as a whole *not* only the percolating cluster. As a matter of fact, all clusters *excluding* the largest one specify the percolation transition in the same way as the percolating cluster does it.

2. Our analysis of realistic cosmological models does not support the conclusion (based on the analysis of toy models) that the dependence of percolation properties on the mean density of galaxies is a serious problem of this method.

3. The strong dependence of percolation properties on the volume of the sample found in some studies was probably caused by an inadequate technique used for measuring percolation thresholds.

We discuss these issues in greater detail below. The rest of the paper is organized as follows. In § 2 we briefly describe the cosmological models which we studied using the percolation technique. Section 3 describes the specifics of our version of the percolation method and the results of the tests. In § 4 we present the analysis of the cosmological models. Finally, § 5 is the summary of our results.

2. COSMOLOGICAL MODELS

Numerical simulations of three cosmological models—the CDM, C + HDM, and $n = -1$—were done using standard Particle-Mesh (PM) code (Hockney & Eastwood 1981; Kates, Kotok, & Klypin 1991).

In C + HDM model the dark matter is a mixture of 30% of hot particles (e.g., tau-neutrinos with the rest mass of 7.2 eV) and 70% of cold particles (Holtzman 1989; Klypin et al 1992). Both the CDM and C + HDM models were normalized to the biasing parameter $b = 1.5$. The $n = -1$ model were normalized to $b = 2$. All models were simulated in a 50 Mpc box ($h = 0.5$). The C + HDM simulation and one of the CDM simulations used the same set of random numbers. The comparison of the two simulations emphasizes the difference between the C + HDM and the CDM models. Note that the CDM and C + HDM models are not very different on the scales within the box: the linear normalization at $r_{\text{top-hat}} = 8\ h^{-1}$ Mpc is the same, and the slope of the power spectra are also quite similar at these scales. So it is really a challenge for a method to detect difference between the models.

The above models were simulated on a 256^3 mesh. The CDM simulations and the $n = -1$ simulation had 128^3 particles each, while the C + HDM simulation has 128^3 cold particles and six sets of hot particles of 128^3 each. The CDM and C + HDM simulations were started at $z = 15$. The power-law simulation, which has much higher amplitude of fluctuations at the smallest resolved scale (195 kpc), was started at redshift $z = 99$. In total three realizations of the CDM model and one realization for each of the C + HDM and $n = -1$ models were simulated. A detailed description and analysis of the CDM and C + HDM models may be found in Klypin et al. (1992). Figure 1 shows positions of about 10% randomly selected particles in a thin (2 Mpc) slice in the C + HDM simulation (left) and in the CDM simulation with the same initial phases (right). The slice was chosen to pass through a dense region near the center of the plot. The magnified inner part of the region where about half of the particles are plotted shows the small-scale difference between the two models: C + HDM model in Figure 2 and CDM model in Figure 3.

We also briefly studied percolation in the C + HDM model simulated in a 100 Mpc box on 512^3 mesh having 256^3 cold particles and 2×256^3 hot particles.

To identify "galaxies" in the model, we simply find all local maxima of the total density above some threshold. Positions of the maxima were treated as the galaxy coordinates. In the C + HDM simulation the number of galaxies was 835, 1824, and 4617 for density thresholds of 100, 50, and 25, correspondingly. In the CDM simulation with the same initial phases the number of galaxies was 751, 1640, and 3318 for density thresholds of 200, 100, and 50, correspondingly. Thus, with the same density threshold the CDM model produced about twice the number of galaxies as the C + HDM model and doubled threshold results in about half the number of galaxies. The $n = -1$ model produced many more galaxies: for the density threshold of 50 it had 5650 galaxies. Thus the number of galaxies in the simulations was not very large and is similar to what one might expect in a volume-limited catalog at the present time or in the near future. Assuming the Schechter luminosity function with parameters $M_* = -19.2$, $\alpha = -1.1$,

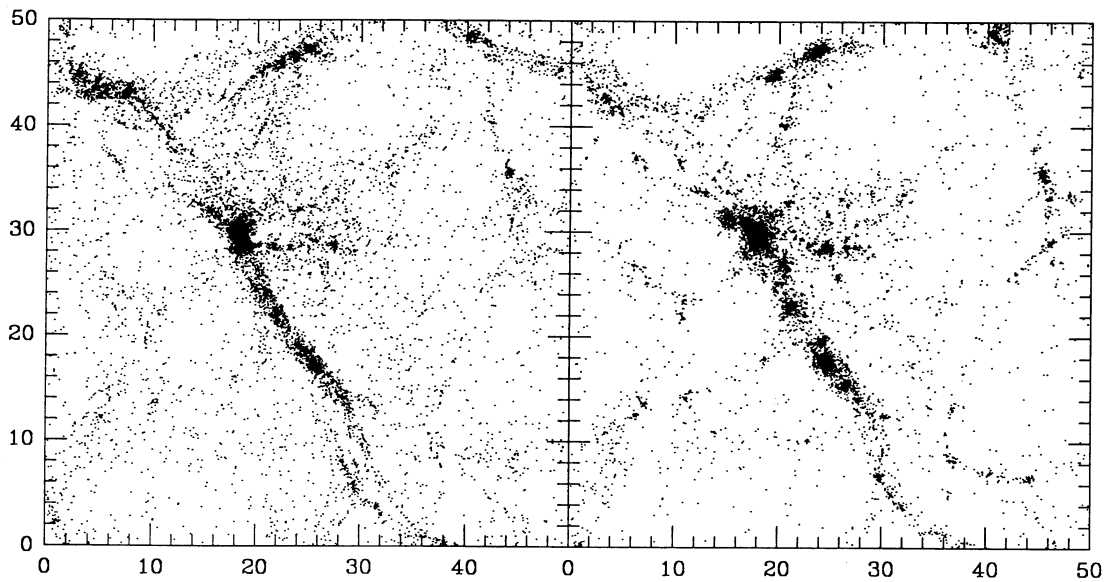

FIG. 1.—Positions of randomly selected particles (about 10% of total) in a thin (2 Mpc) slice in the C + HDM simulation (*left*) and in the CDM simulation with the same initial phases (*right*). Coordinates are scaled in Mpc ($h = 0.5$).

and $\phi_* = 0.02\ h^3\ \mathrm{Mpc}^{-3}$ (de Lapparent, Geller, & Huchra 1989) one may expect to find about a thousand galaxies 3.2 magnitude weaker than M_* in a cube like ours. The low limit of the maximum density imposes a limit on the mass of galaxies. For example, the density threshold 50 imposes the restriction on the mass of galaxies $M > 2.5 \times 10^{10}\ M_\odot$.

3. PERCOLATION TECHNIQUE

In this paper we use percolation analysis on three-dimensional cubic lattice assuming that each cell can be immediately linked to no more than the closest six neighbors. In percolation theory jargon, this is called site percolation on a simple cubic lattice. It is worth stressing that in recent studies (de Lapparent et al. 1991; Mo & Börner 1990) different schemes were used, allowing more connections, which of course makes percolation happen easier that is, at lower filling factors. Unfortunately it also reduces the percolation threshold for the Poisson distribution, which means that the signal-to-noise ratio does not improve. An advantage of our definition of the neighborhood is that it has been very well studied. Before we present the results of our tests let us describe the phenomenon of percolation in greater detail.

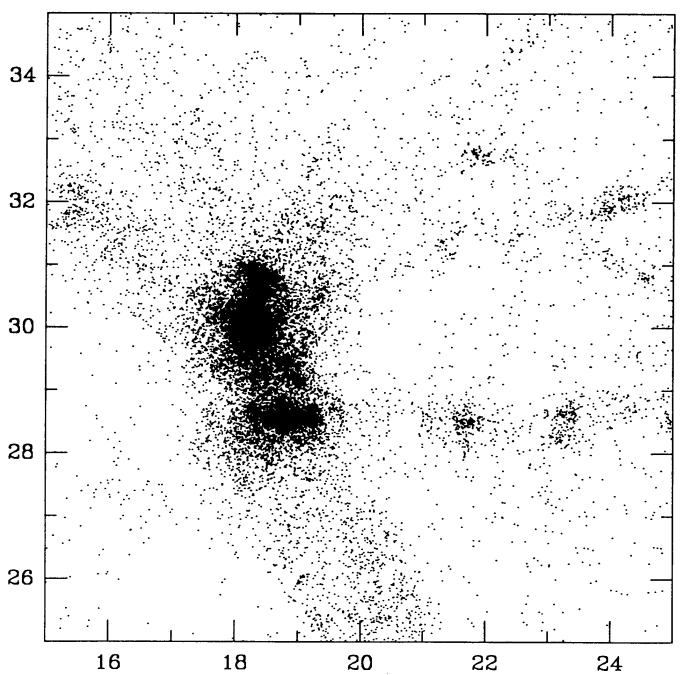

FIG. 2.—Distribution of cold particles in the C + HDM model in 10 Mpc × 10 Mpc × 2 Mpc region close to the center of Fig. 1. The number of particles in 13,680, which is a half of the total number in the region.

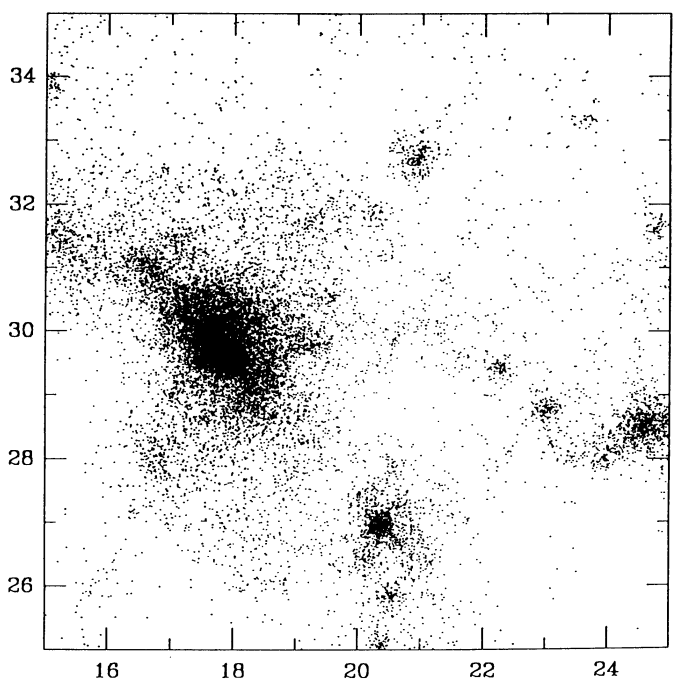

FIG. 3.—Same as in Fig. 2, but for the CDM simulation. The number of particles is 13,185.

As we stressed in the Introduction, the most essential concept of percolation theory (in infinite systems) is the critical transition from the state where there are only finite clusters, to the state where an infinite cluster exists. Probably the simplest way to illustrate the formation of an infinite cluster is to observe the growth of the diameter of the largest cluster in a system: at the transition it becomes infinite. In an ealier paper by one of the authors (Shandarin 1983) the growth of the largest cluster was used mainly for an illustrative purpose, rather than a suggestion that it is the only or the best method of detecting percolation in a system. Unfortunately, in some later papers (e.g., Bhavsar & Barrow 1983; Dekel & West 1985) this property of the percolation phenomena was treated almost as a synonym for percolation and occasionally was used for estimating the percolation threshold in samples which are not well suited to this technique (e.g., a cone). That was part of the reason for the reported problems in applying the percolation statistics of cosmology.

The formation of a cluster spanning the entire volume (which we call the "direct test for percolation") is only one manifestation of the critical behavior of a system near the percolation transition. Although quite intuitive, it has serious disadvantages when used as an estimator of the percolation thresholds. First, it is not clear how to apply the direct test to samples having shapes other than a cube or parallelepiped, for example, a sphere or a cone. Second, the direct test is very sensitive to what happens near the boundaries: a fluctuation or an incompleteness of the catalog could significantly affect the results. The diameter of the largest cluster as a measure of the percolation threshold is free of the first disadvantage but suffers from the second. Fortunately they are not the only features of percolation. There are others which much better suit the purpose of measuring the percolation thresholds. One is related to the growth of the size (the size equals the volume equals the number of cells) of the largest cluster and another to the weighted mean size of all clusters, excluding the largest one. It is clear that the growth of the diameter of the largest cluster is strongly suppressed when its end points reach the boundaries of the region. On the contrary, because the volume can grow by joining the cells along its whole surface, most of which are perfectly inside the region, it is less sensitive to boundary effects. Using both parameters in cosmology was suggested by Klypin (1987), and the size of the largest cluster was used as an estimator of the percolation threshold by de Lapparent et al. (1991).

When describing the percolation transition, it is convenient to introduce the multiplicity function $n(v)$ which is the average numer (per lattice site) of clusters of size v. The multiplicity function depends on the fraction of filled cells p, which we call the filling factor . In general, every cell has one of three options:

1. It can be empty with the probability $1 - p$;
2. It can be a member of the infinite cluster with the probability μ_∞; or
3. It can be a member of a finite cluster with the probability $\sum_v v \times n(v)$.

The sum of all probabilities is of course unity:

$$1 - p + \mu_\infty + \sum_v v \times n(v) = 1 \,. \tag{1}$$

At the percolation threshold the whole system experiences a dramatic transition. For instance, at the percolation verge p_c the multiplicity function is simply a power law (Stauffer 1979, 1985); $n(v; p_c) \propto v^{-\tau}$, $\tau \approx 2.1$. But it is not true in the general case where it falls exponentially at high v. In principle, the form of the multiplicity function could be used to detect the critical point. However, keeping in mind the limited amount of observational data, it seems more reasonable to begin with an analysis of the lowest moments of the multiplicity function, which also exhibit singular behavior. The singular terms for the lowest moments are as follows:

$$\begin{aligned} \sum_v n(v) &\propto |p - p_c|^{2-\alpha} \,, \\ \sum_v v \times n(v) &\propto (p - p_c)^\beta \,, \\ \sum_v v^2 \times n(v) &\propto |p - p_c|^{-\gamma} \,, \end{aligned} \tag{2}$$

where α, β, γ are critical exponents (e.g., Stanley 1971). It also can be shown from equations (1) and (2) that $\mu_\infty \propto (p - p_c)^\beta$.

In this paper we deal mostly with two parameters: the fraction of volume occupied by the largest cluster, normalized to one cell,

$$\mu_\infty = v_{\max}/N_t \,, \tag{3}$$

and the mean weighted size of the clusters excluding the largest one:

$$\mu^2 = \frac{\sum_v v^2 \times n(v)}{N_t^{2/3} \sum_n n(v)} \,, \tag{4}$$

where $N_t = N_x \times N_y \times N_z$ is the size of the lattice, and the summation is over all clusters excluding the largest one. We have introduced the additional factor $N_t^{2/3}$ in the denominator (compared to the standard defintion of μ^2 in theory of percolation) to make plotting of the results easier. According to percolation theory, the critical exponents are $\beta \approx 0.4$ and $\gamma \approx 1.7$ (Stauffer 1985). We expect that μ_∞ grows rapidly and μ^2 has a sharp peak in the vicinity of the percolation threshold.

3.1. *Poisson Lattice*

A Poisson distribution of cells with the probability p for a cell to be filled is a simple but quite interesting model. In spite of its simplicity, the growth of clusters in the model provides a reasonably accurate model for a number of critical phenomena (Stanley 1971). We show in § 4 that in the vicinity of the percolation threshold the growth of largest cluster in both the CDM and the C + HDM models is similar to that in the Poisson model. We use the Poisson model for the purpose of testing the code and studying the dependence of percolation properties on the size of the lattice and the shape of the volume.

The dependence of the percolation parameters on the size of the cubic lattice is shown in Figure 4 for 33^3, 63^3, 127^3, and 256^3 lattices. One hundred different sets of random numbers were averaged for each lattice except the 256^3 lattice, for which 40 sets were used in the vicinity of the percolation transition (three closest points) and 10 sets for other filling factors. The top panel shows the fraction of percolating realizations for a given filling factor p. This is the direct test for percolation. The middle panel shows the fraction of cells in the largest cluster μ_∞ (eq. [3]) as a function of the filling factor. Here μ_∞ approaches zero at $p < p_c$ and the asymptote $\mu_\infty \propto (p - p_c)^{0.4}$ at $p > p_c$. The bottom panel shows the dependence of the weighted mean volume of the remaining clusters μ^2 (eq. [4]). With the growth

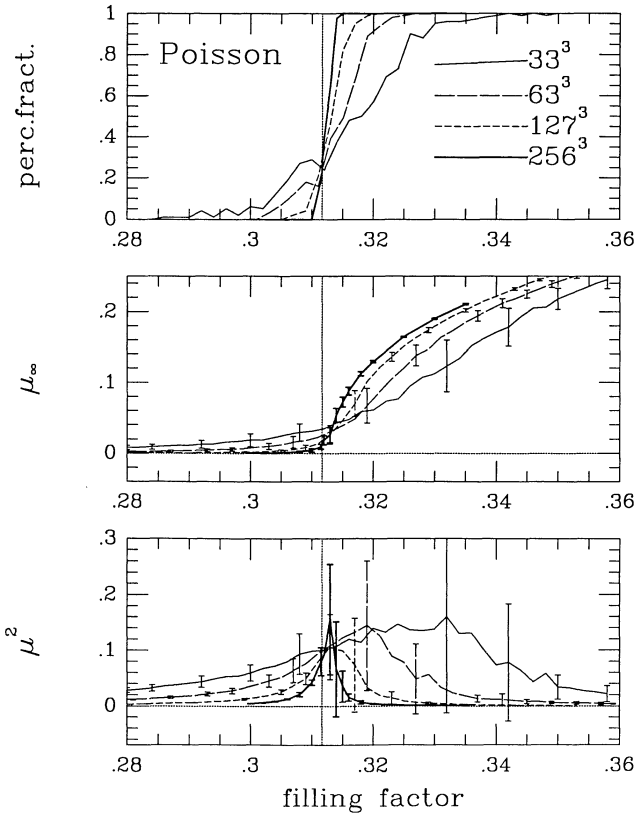

FIG. 4.—Percolation in the Poisson distribution on the lattices from 33^3 to 256^3 is shown. The error bars correspond to 1 σ dispersion. The top panel shows the fraction of percolating systems for given filling factor p (direct test for percolation). The middle panel shows the fraction of cells in the largest cluster, μ_∞. The bottom panel shows the dependence of the mean square of the volume of clusters, μ^2, excluding the largest one. The vertical dotted line marks the percolation threshold on very large lattices.

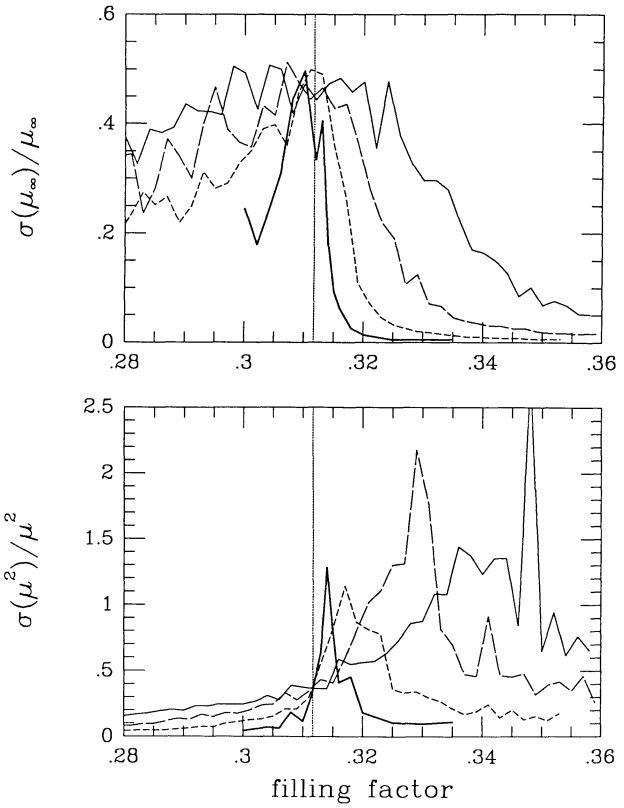

FIG. 5.—Relative level of fluctuations for two estimates of the percolation threshold μ_∞ (*top panel*) and μ^2 (*bottom panel*) shown in Fig. 4. The vertical dotted lines marks the percolation threshold on very large lattices. The types of lines are as in Fig. 4.

of the lattice size, the maximum of $\mu^2(p)$ becomes narrower and its position approaches the percolation threshold p_c. The vertical dashed line marks the percolation threshold measured on very large lattices (e.g., Stauffer 1985). The central panel shows that the transition to the percolating phase is quite sharp on all lattices. Even using the smallest 33^3 lattice, one can estimate the percolation threshold with a relative accuracy better than 5% which corresponds to the range $0.296 < p_c < 0.327$. The bottom panel shows that the μ^2 statistics is somewhat noisier for small lattices. We show the dispersions of both μ_∞ and μ^2 in units of corresponding mean values in Figure 5. One can see that the noise in μ_∞ is relatively large in the nonpercolating regime even for our largest lattice and falls rapidly in the percolating regime. The μ^2 statistics shows the opposite behavior; however, the noise is generally higher. In a sense, the μ_∞ and μ^2 statistics are complementary.

3.2. *Anisotropic Lattices*

In applying the percolation technique to real galaxy catalogs, there is the problem of inhomogeneity of the catalogs caused by the dependence of the selection function on the distance. One way of dealing with this problem is the introduction of correction for the effect of the selection function. In the context of the percolation analysis, it was used in two slightly different forms by Bhavsar & Barrow (1983) and de Lapparent et al. (1991). Bhavsar & Barrow, who studied percolation in pointlike systems, used the neighborhood radius varying with the distance, and de Lapparent et al. proposed using a lattice with cells of varying sizes. The volume of both a sphere and a cell is approximately inversely proportional to the mean density of galaxies. Although somewhat technically different, both methods suffer from a common disadvantage. To see it clearly let us discuss the following *gedanken* experiment. Let us imagine that we have a very large volume-limited sample of galaxies. For the purpose of testing the percolation technique we produce a subsample by simulating the selection of galaxies with a specific selection function. Now if we find the percolation thresholds separately for two parts of the subsample: one in a nearby region where the subsample is complete and the other in a remote region where the subsample becomes a Poisson distribution, they may be very different despite that they were similar in the original sample. Assuming that the original distribution is of a filamentary type, the percolation threshold in the nearby region is small. In the remote region it may be equal to the Poisson value. Combining them together, we may end up in the mixture dominated by the Poisson part. The suggested corrections (the variable radius of spheres, or size of cells) does not help at all because it cannot restore the lost information and does not reduce the contribution of the Poisson part.

We can suggest a different approach which does not mix close and remote regions. Let us divide the volume of a sample into layers perpendicular to the line of sight. The thickness of a

layer must be chosen to keep the density roughly constant within the layer, and the layers must be analyzed separately. Obviously free of the disadvantage of the first technique it can be affected by the anisotropy of the layers. In order to satisfy the requirement of the homogeneity, at least some of the layers must be relatively thin. In particular, one can expect that on lattices having the shape of a flattened square parallelepipeds the percolation threshold must approach the two-dimensional value in the limit of small thickness. The fully study of this effect must produce the percolation threshold as a function of distance for each model and a specified selection function. And this function must be compared with a similar function calculated from a real galaxy catalog.

Here we only make a rough estimate of the significance of this effect. In Figure 6, we show the dependence of the percolation threshold on the thickness of the lattice. As an sample we show percolation properties on a series of lattices having the same square base of 127^2 and the thickness varying from 1 to 8; we also show the percolation parameters for the full cubic lattice 127^3. In the anisotropic regions, the direct percolation test was applied only along the two largest sides. The percolation parameters were evaluated for 100 realizations of random numbers of each lattice. As expected, the percolation parameter decreased from the two-dimensional value ($p_c = 0.57$) when the lattice had only one layer of cubic cells to almost the three-dimensional value when it had thickness 8 (roughly 0.34–0.35 instead of 0.31). Similar calculations for the 33^3 lattice show very similar results, but with greater dispersions due to smaller statistics. In the structureless Poisson distributions, the percolation thresholds depend only on *absolute* thickness of the lattice.

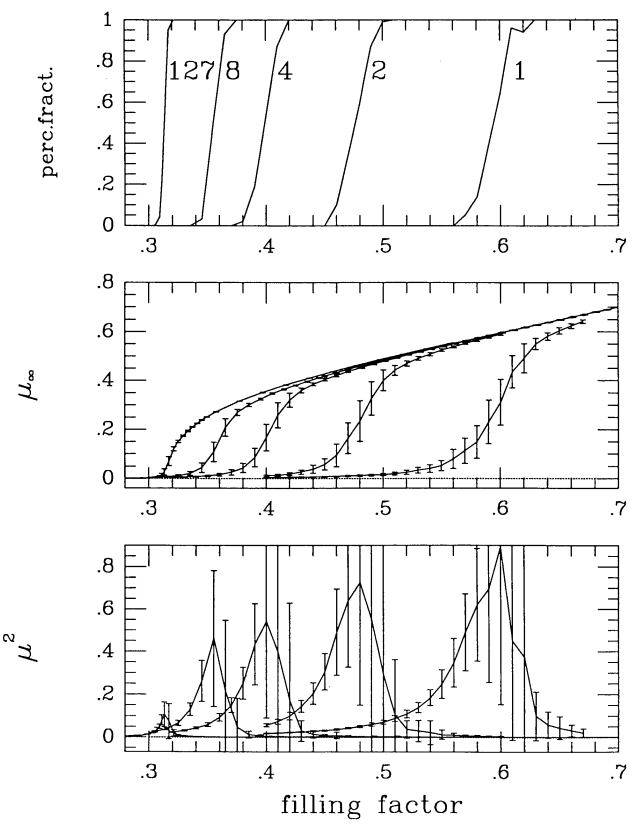

FIG. 6.—Dependence of the percolation thresholds on the thickness of the parallelpiped given in mesh units; 1 corresponds to the two-dimensional lattice.

3.3. *Discrete Distributions*

When we dealt with the Poisson models, we could increase the filling factor until percolation occurred. Studying percolation in mass distribution one can vary the filling factor by changing the density threshold when labeling the cells. Galaxy samples are discrete distributions and usually do not percolate without smoothing. One can think of a number of smoothing filters: the Gaussian smoothing (Mo & Börner 1990; Dominik & Shandarin 1992), the variation of the sizes of cells (Einasto et al. 1984; de Lapparent et al. 1991), a construction of spheres of varying radius about each galaxy (Shandarin 1983; Bhavsar & Barrow 1983; Dekel & West 1985; Klypin 1987).

Here we suggest a very simple smoothing technique, which, in addition to being very efficient computationally, also roughly preseves the Poisson characteristics when applied to a Poisson sparse distribution. Applying the smoothing once means that (1) every originally filled cell remains filled, and (2) each originally empty cell changes its type if and only if it has at least one closest (of six total) neighbor being originally filled. Applied sequentially this procedure eventually fills the gaps between galaxies and a percolating cluster forms. Roughly speaking the procedure blows a cube around each galaxy with the size measured in cell units equal to the number of times the procedure has been applied.

To study percolation in our galaxy samples we must be able to start from very low filling factors. As mentioned above in our simulation we selected only a few thousand galaxies which corresponds to a filling factor $p < 3 \times 10^{-3}$ on the 256^3 lattice. Testing this technique, we randomly labeled a few thousand cells as filled on the 256^3 lattice and applied the smoothing procedure until percolation occurred. We found that the percolation threshold in such a system is about $p_c = 0.28$, compared to the $p_c = 0.31$ in real Poisson distributions. It is worth mentioning that the Gaussian filter tends to transform this distribution to the Gaussian one having a percolation threshold about $p_c = 0.16$ (Dominik & Shandarin 1992). Another disadvantage of the Gaussian filter is that it brings about an additional parameter which is the density threshold. The major disadvantage of our method is that it cannot increase the filling factor continuously.

4. PERCOLATION IN COSMOLOGICAL MODELS

4.1. *Dark Matter*

Three cosmological models were simulated to study connected structures: the CDM, C + HDM, and the model with the power-law initial spectrum ($n = -1$). In all models the percolation properties of the mass distribution were studied on a 256^3 lattice. The density threshold $\rho_{\rm thr}$ served as a free parameter. Cells with density above $\rho_{\rm thr}$ were labeled as filled and all the rest as empty. Thus, the filling fraction p was a function of the threshold density $\rho_{\rm thr}$. The size of the largest cluster μ_∞ and the mean size of all the rest of the clusters but the largest one μ^2 averaged over three realizations of the CDM model are shown in Figure 7. The two panels show the typical features of percolation: the fast growth of the largest cluster and a peak in $\mu^2(p)$. Similarly to the Poisson distribution, the fluctuations in the critical region are quite large (especially for μ^2), but the critical transition can be easily identified some-

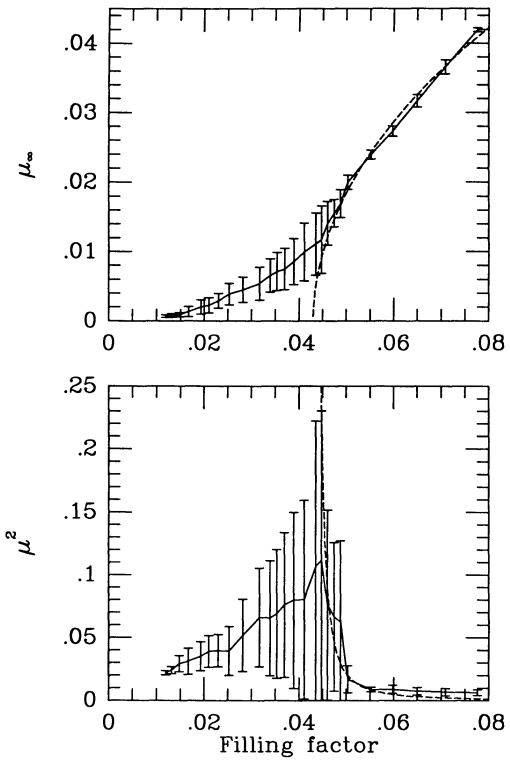

FIG. 7.—Percolation in the CDM model, averaged over three realizations. The dashed curve in the top panel is the fit $\mu_\infty = 0.22(p - p_c)^{0.5}$, the dashed curve in the bottom panel is $\mu^2 = 0.3(p - p_c)^{-1.7}$. In both cases $p_c = 0.043$.

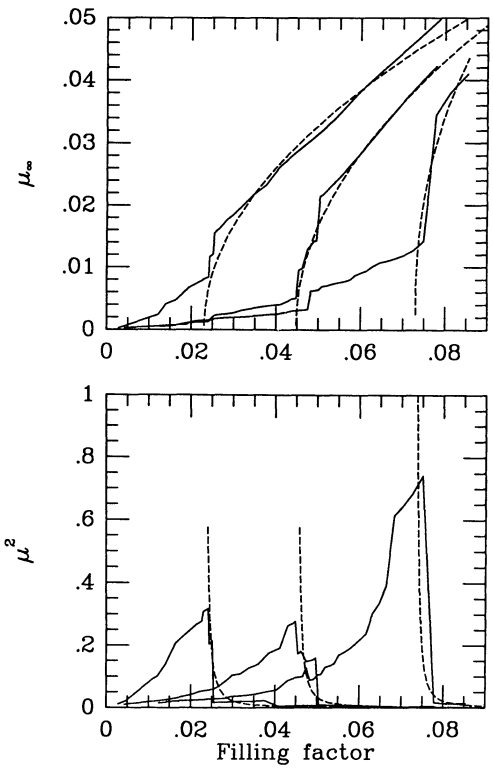

FIG. 8.—Percolation properties of the C + HDM, CDM, and $n = -1$ models are shown correspondingly from left to right. The C + HDM and CDM models have the same set of random numbers. The dashed curves in the top panel are the fits $\mu_\infty = A(p - p_c)^\beta$ with $p_c = 0.0230$, 0.0448, 0.073, $A = 0.20$, 0.23, 0.25, and $\beta = 0.5$, 0.5, 0.4 for the C + HDM, CDM and $n = -1$ models respectively. The dashed curves in the bottom panel are the fits $\mu^2 = 0.3(p - p_c)^{-1.7}$ with the same percolation thresholds as above.

where between $p = 0.04$ and $p = 0.05$. The percolation threshold can be evaluated more accurately by fitting the data points by equation (2). Both $\mu_\infty(p)$ and $\mu^2(p)$ can be reasonably well approximated above the percolation threshold as $\mu_\infty(p) = 0.2(p - p_c)^{0.5}$ and $\mu^2(p) = 0.3(p - p_c)^{-1.7}$. We found that the exponent 0.5 approximates μ_∞ slightly better than the Poisson value 0.4, but the difference is small, and we do not claim that it is significant. This gives the percolation threshold $p_c = 0.043 \pm 0.005$ for the dark matter in the CDM model normalized to the biasing parameter $b = 1.5$. Results of the direct test (the detection when the largest cluster links the opposite sides of the cube) are consistent with this estimate. We also roughly estimated the uncertainty in the percolation threshold from variations of the value from one sample to another and by comparing various statistics (the direct test, μ_∞, μ^2). For the C + HDM simulation the same procedure gives the threshold $p_c = 0.0230 \pm 0.005$, which is significantly lower than that in the CDM model. Figure 8 compares the CDM (middle curves) and C + HDM (left curves) simulations having the same set of initial random numbers. The results for the $n = -1$ model are also shown (right curves). The growth of the largest cluster is approximated as $A(p - p_c)^{0.5}$ (dashed curves) with the parameters $A_{\rm CDM} = 0.23$, $A_{\rm C+HDM} = 0.20$, $p_{c,\rm CDM} = 0.0448$, and $p_{c,\rm C+HDM} = 0.0230$. For the $n = -1$ model the approximation is $\mu_\infty = 0.25(p - p_c)^{0.4}$, $p_{c,n=-1} = 0.073$. The dashed curves in the bottom panel are $\mu^2 = 0.3(p - p_c)^{-1.7}$ with the same percolation thresholds as above.

We found that in the vicinity of the critical point, the growth of the largest cluster can be approximated by roughly the same exponent, $\beta = 0.4$–0.5, independently of the model. The only difference is the amplitude. To show this we have plotted μ_∞ as a function of $p - p_c$ on logarithmic scale. Figure 9 shows log (μ_∞) for the CDM model (averaged over three realizations) and the C + HDM simulation as well as for the Poisson distribution. The size of the largest cluster μ_∞ for the Poisson distribution was scaled down by a factor of 6.7. All three models predict almost the same relative rate of growth of the largest cluster. The only difference is that the volume of the largest cluster in the Poisson distribution is considerably larger than the C + HDM and CDM simulations.

In order to distinguish between filamentary and cellular structures, we made the percolation analysis of empty cells which were the cells having the density *below* the threshold. We found that even for densities as low as $\rho < 0.1\langle\rho\rangle$ there was percolation through empty cells. As a matter of fact, in the case of the CDM model there was a single void occupying most of the volume (85%) and only a few very small voids. We could not reach lower densities because of discreteness. Thus, high-density walls do not isolate voids and therefore do not form a cellular structure in our models. Of course, it does not exclude the existence of isolated walls (pancakes) in the distribution of the dark matter. We also cannot exclude that a cellular structure forms at even lower density thresholds that would be in agreement with the results obtained for two-dimensional distributions (Dominik & Shandarin 1992).

One can use the concept of the percolation analysis to formulate a quantitative topological criterion of a cellular structure. A structure can be defined to be cellular if the percolation

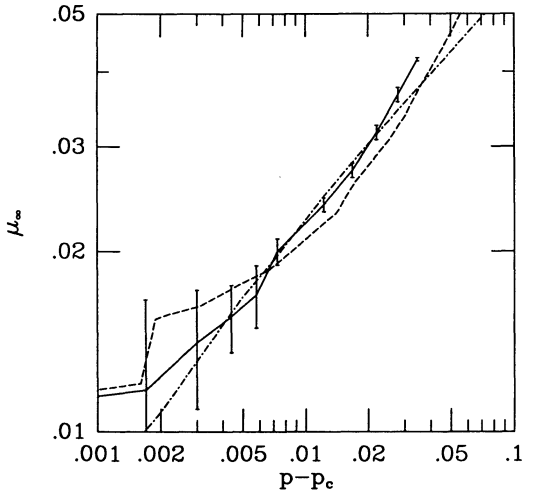

FIG. 9.—Growth of the volume of the largest cluster in the vicinity of the critical point: the CDM model (*solid line with error bars*), the C + HDM model (*dashed curve*), the Poisson distribution (*dot-dashed curve*). The Poisson curve is scaled down by a factor of 6.7.

threshold through empty cells is smaller than $p_c < 0.84$ for a continuous distribution or $p_c < 0.69$ for a discrete distribution. Here the p_c is the filling factor of *filled* cells. Therefore this definition requires that percolation through the *empty* cells is suppressed when the fraction of empty cells is large: $p_c(\text{empty}) > 0.16 = 1 - 0.84$ for continuous distributions or $p_c(\text{empty}) > 0.31 = -0.69$ in discrete distributions. This definition emphasizes the topological aspect of a cellular distribution and may include the case when the thickness of walls is not much less than the diameters of voids.

The above percolation thresholds characterize the stage of the evolution identified with the present time. However, the percolation thresholds change with time. At the initial time ($z = 15$ for the CDM and the C + HDM, or $z = 99$ for the $n = -1$ model) the percolation threshold was $p_{c,\text{linear}} = 0.16$–$0.18$, which is close to the percolation threshold for Gaussian fields and may be identified with the percolation level of density fluctuations in the linear regime. Later the threshold decreases in all models indicating the formation of elongated anisotropic structures. In the $n = -1$ model the threshold had a minimum $p_{c,n=-1} = 0.04$ at the redshift $z \approx 2$ and grew up to the level 0.073 by the present time. The CDM and the C + HDM simulations also show a minimum in $p_c(z)$ but at about $z = 0.3$ ($p_{c,\text{CDM}} \approx 0.035$, $p_{c,\text{C+HDM}} \approx 0.022$). The minimum is quite shallow: in the CDM model the value similar to the threshold at $z = 0$ was also at $z = 2.7$.

The threshold density at the percolation ρ_c also varies with time. In the linear regime the density threshold was slightly greater than twice the mean density ($\rho_c = 2.2$–2.4). As the fluctuations enter the nonlinear regime, the density ρ_c increases. Approximately at the same time, when $p_c(z)$ has a minimum, the density ρ_c had a maximum which was different in different models: $\rho_{c,\text{C+HDM}} \approx 6.0\langle\rho\rangle$, $\rho_{c,\text{CDM}} \approx 5.5\langle\rho\rangle$, $\rho_{c,n=-1} \approx 3.1\langle\rho\rangle$. After passing the minimum, the density threshold decreases. At $z = 0$ it is practically the same in the C + HDM model $\rho_{c,\text{C+HDM}} \approx 6.0\langle\rho\rangle$; however, it is significantly lower in other models: $\rho_{c,\text{CDM}} \approx 3.3\langle\rho\rangle$, $\rho_{c,n=-1} \approx 0.25\langle\rho\rangle$.

Percolation in the mass distribution in the C + HDM model simulated in a 100 Mpc box is illustrated in Figure 10. The simulation has the same spatial resolution as in the 50 Mpc box, but occupies a larger volume. The percolation threshold for the simulation was $p_c = 0.008$–0.015, $\rho_c = (8$–$12)\langle\rho\rangle$. In this simulation we studied percolation by splitting the box into four or eight equal flattened parallelepipeds in addition to percolation in the full cube. Comparing two statistics we estimate the percolation threshold $p_c = 0.008$–0.026 with a slight indication that it is systematically higher in thinner parallelepipeds, which is not unexpected.

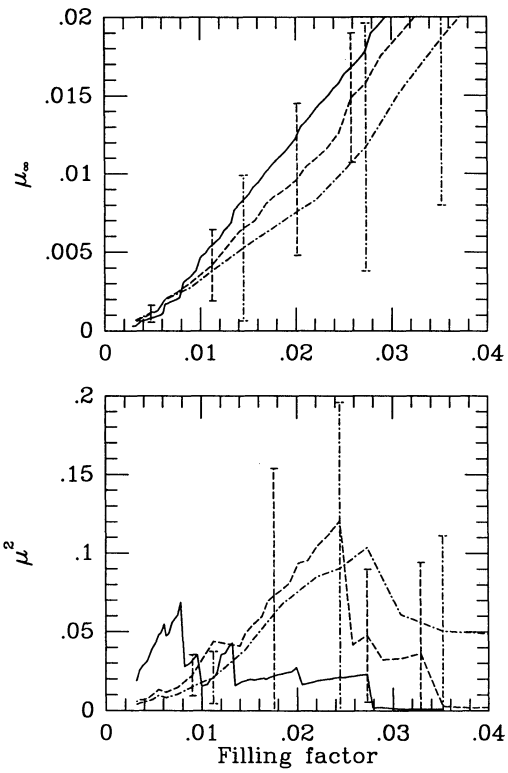

FIG. 10.—Percolation in the mass distribution in the C + HDM model simulated in a 100 Mpc box. Solid lines show the percolation parameters for the full cube (512^3). Dashed lines with error bars are the parameters averaged over four slices of 128 cells thick; and dot-dashed lines are averaged over eight slices of 64 cells thick.

4.2. *Galaxies*

Dark halos, which were identified with galaxies, are not distributed identically to the dark matter. This was indicated by the analysis of the correlation functions and velocities (Klypin et al. 1992). Although dark halos are more strongly clustered, even on large scales, it does not necessarily mean that the distribution of galaxies (halos) is more filamentary than that of the dark matter. The density threshold ρ_c at percolation was not very high: about 3–6 in the CDM and C + HDM models, and even less than unity in the $n = -1$ model. This may mean that there are relatively low density dark matter "bridges," which link small numerous clumps to the largest ones and which actually account for the fast growth of the percolating cluster in the mass distribution. However, many of these "bridges" do not have galaxies. As a result, in the galaxy distribution the percolation may happen at larger filling factors. There is also another cause for the difference in percolation properties in the galaxy and dark matter distributions. It is the noise due to the discreteness of galaxy distributions. As we mentioned above, a distribution of pointlike objects

(galaxies) must be somehow smoothed before clusters can be formed. No matter how the smoothing is done (say, by using filters on lattices, as we do, or by surrounding each galaxy with a sphere as in the original percolation prescription), it results in thicker structures and greater filling factors at percolation. The effect is inevitable—this is the price for the lost information on small scales.

Dekel & West (1985) were concerned that the percolation threshold scales with the mean number of objects in a way which depends on the dimensions of the underlying set. Although their arguments were based mainly on the analysis of overly simplistic toy models and they used different methods for measuring percolation thresholds, we studied this effect in our models.

Figure 11 shows the results of the percolation analysis of the distribution of galaxies identified with the maxima of density higher than $50\langle\rho\rangle$. The constant threshold results in different numbers of objects in different simulations. The difference between the largest sample in the $n = -1$ model and the smallest one in the C + HDM model is about a factor of 3. Comparing the percolation thresholds obtained from the μ^2 and μ_∞ statistics, we estimate p_c as follows: $p_{c,\mathrm{C+HDM}} = 0.08$, $p_{c,\mathrm{CDM}} = 0.11$, $p_{c,n=-1} = 0.15$ with the uncertainty about ± 0.02. The thresholds were estimated by fitting the data with equation (2). Although the difference between the CDM and C + HDM models is not large, Figure 11 demonstrates that the percolation analysis does distinguish the models. The same smoothing filter was applied to 3200 randomly distributed particles, which was close to the number of galaxies in the CDM simulations. On the one hand, the percolation threshold of the sparse Poisson distribution ($p_c = 0.28$–0.32) is not much different from the known value ($p_c = 0.31$). But on the other hand, it is significantly greater than in our cosmological models. We conclude that all three cosmological models are very filamentary as well as that there is a significant difference in the percolation properties between them.

The dependence of the percolation thresholds on the number and the mass of galaxies is shown in Figure 12. the density thresholds identifying galaxies were 50, 100, 200 in the CDM model and 25, 50, 100 in the C + HDM simulation. There is some indication that the curves may shift to the smaller thresholds as the number of objects grows, but the effect is obviously quite weak. Figure 13 illustrates percolation in different models when the number of galaxies is kept almost the same ($N_g = 4150$–4250). We found the following thresholds: $p_{c,\mathrm{C+HDM}} = 0.06$, $p_{c,\mathrm{CDM}} = 0.10$, and $p_{c,n=-1} = 0.16$. The difference between CDM and C + HDM models now looks a little more distinct which is in a rough agreement with the previous result. In this figure the number of galaxies in the C + HDM model is 2.3 times greater than that in Figure 11. Therefore, less smoothing is needed to reach the percolation level of the filling factor. We conclude that although the effect discussed by Dekel & West (1985) strictly speaking exists, it does not seem to be very important in realistic models.

The analysis of voids in the galaxy distributions gave a result similar to that for the mass distribution. Applying our smooth-

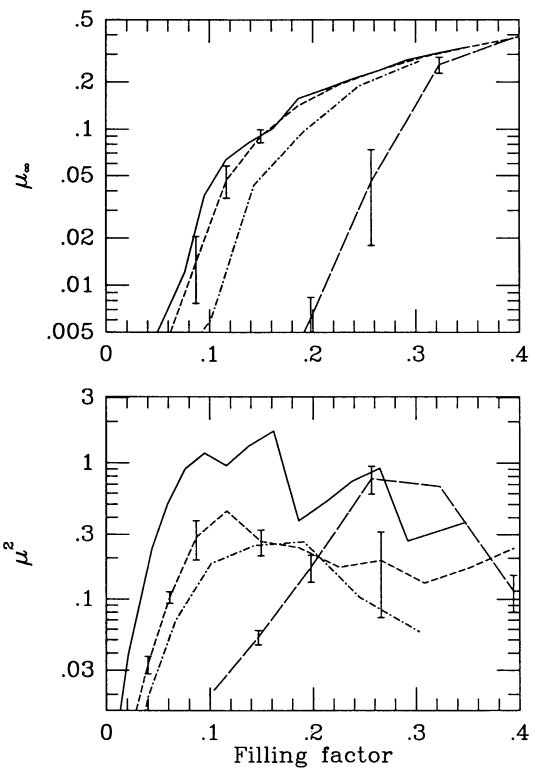

FIG. 11.—Percolation in galaxy distributions is shown: the C + HDM model (*solid line*); three CDM simulations (*dashed curve with error bars*); the $n = -1$ model (*dot-dashed curve*); and the Poisson distribution of 3200 points (*long dashed curve with error bars*). The density threshold $50\langle\rho\rangle$ or the mass limit $2.5 \times 10^{10}\ M_\odot$ was used to identify galaxies in every model.

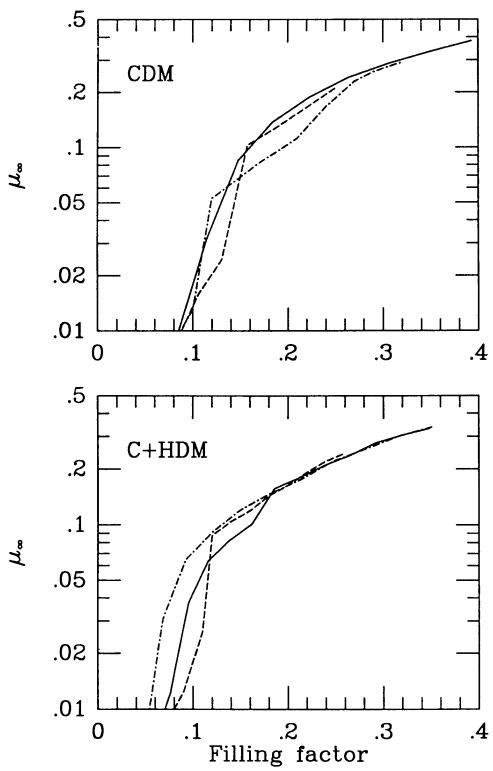

FIG. 12.—Volume of the largest cluster for the galaxy distributions specified by different masses and therefore having different numbers of galaxies. Both the CDM and the C + HDM simulations had the same initial phases of fluctuations. For both models the solid curves correspond to $M > 2.5 \times 10^{10}$ $M_\odot$ ($N_{\mathrm{CDM}} = 3318$, $N_{\mathrm{C+HDM}} = 1824$); the dashed curves correspond to $M > 5 \times 10^{10}$ $M_\odot$ ($N_{\mathrm{CDM}} = 1640$, $N_{\mathrm{C+HDM}} = 835$); and the dashed-dotted curves correspond to $M > 1.25 \times 10^{10}$ $M_\odot$, ($N_{\mathrm{C+HDM}} = 4617$), and $M > 10^{11}$ $M_\odot$ ($N_{\mathrm{CDM}} = 751$).

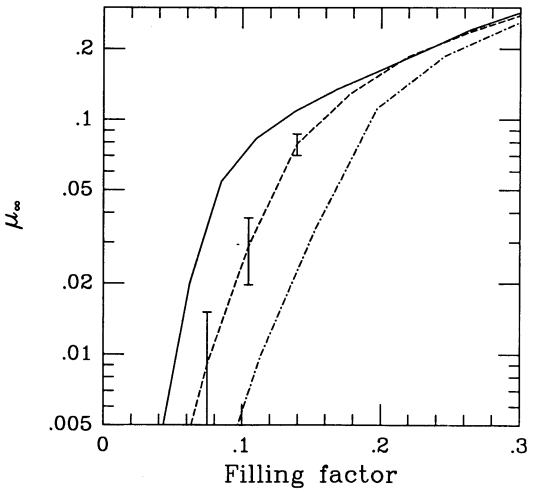

FIG. 13.—Volume of the largest cluster when the number of galaxies was about the same. The solid curve shows the C + HDM simulation; the dashed curve with error bars is for the CDM model (averaged over three realizations); and the dot-dashed curve is for the $n = -1$ simulation. The number of galaxies is roughly the same: $N = 4150$–4250.

ing filter (see § 2) 20 times (thus, each isolated galaxy was blown up to a cube of 40 cells in size), the stage when the fraction of the volume in voids was only 40%, there was still a percolating void, encompassed most of the empty cells. We conclude that the galaxy distributions in the models in question also do not look like thin walls separating large voids. Thus, mathematical models, assuming that the galaxy distributions looks like thin walls separating voids (e.g., Voronoi tesselation), are not very good approximations for the CDM or C + HDM cosmological models. The CDM and C + HDM models, rather, assume the existence of a system of thin filaments, connected in "knots" where clusters of galaxies or groups are located. However, walls (pancakes) can also be present.

5. SUMMARY

We have developed a very fast numerical code allowing the detection of percolation on three-dimensional cubic lattices in parallelepiped regions, along with cluster analysis of both the density and point/galaxy distributions. The code is based on the algorithm published in Stauffer (1985). In the case of point distributions, we suggested and tested a new very simple and efficient smoothing technique allowing the study of percolation in point systems using our percolation code.

In order to evaluate the percolation thresholds, we calculated two parameters characterizing the percolation transition along with the direct detection of percolation. One of them is μ_∞, the size (= volume) normalized to the size of the lattice (eq. [3]), and the other μ^2, the weighted mean size of all clusters excluding the largest one (eq. [4]). We found that both can be used for the purpose of evaluating the percolation thresholds, but for small lattices (less than roughly 64^3) the former is somewhat less noisy. A very significant advantage of these two parameters over the direct detection of percolation or the diameter of the largest cluster is that they are much less sensitive to the boundary effects.

As expected, we found that the percolation threshold in square parallelepipeds ($N^2 \times M$) increases with decreasing M: for example, it approaches the two-dimensional threshold when $M \to 1$ for the Poisson distributions. However, in the Poisson distributions, which have no scale, for $M = 8$ the percolation threshold is only about 10% greater than in cubic regions independent of N (except for the level of noise) in the range $33 \leq N \leq 127$. The change in percolation properties through the transition from three-dimensional to two-dimensional lattices is regular and can be easily taken into account if asymmetric regions needed to be studied. This effect is definitely present in the analysis of de Lapparent et al. (1991) (at least when they studied the slices separately), resulting in the increase of the percolation threshold. On the other hand, using the scheme allowing connection with 26 neighbors (compared to 6 in the standard percolation analysis) they definitely decreased the percolation threshold. Even the sign of the net result is difficult to assess without additional analysis.

We studied percolation properties of three cosmological models: CDM, C + HDM, and the power-law model with the slope $n = -1$ in the $\Omega = 1$ universe using N-body simulations. Both the mass and galaxy distributions were studied. We found that for all models the percolation thresholds evaluated by different methods (direct percolation, the largest cluster, and the mean size of all clusters except the largest one) give consistent results.

The percolation thresholds of the mass distributions are significantly different in the three models. As expected, the percolation threshold in the C + HDM is the smallest $p_c = 0.023$; in the CDM model $p_c = 0.044$; and in the $n = -1$ model $p_c = 0.073$. Our estimate of uncertainty of the thresholds is ± 0.005. All thresholds are significantly smaller than the Gaussian value $p_c = 0.16$. Therefore we conclude that as concerns the mass distribution the percolation statistics measures the degree of filamentarity and is able to distinguish between different models.

The percolation thresholds in the simulated galaxy distributions are significantly greater than those in the mass distributions: $p_c = 0.06 \pm 0.02$ for the C + HDM simulation, $p_c = 0.10 \pm 0.02$ for the CDM simulations, and $p_c = 0.15 \pm 0.02$ for the $n = -1$ model. This is not surprising because the random component is much higher due to a small number of galaxies in the samples: from about 800–5000 on the 256^3 lattice. However, the percolation threshold is much less than $p_c = 0.28$ corresponding to the sparse Poisson distribution of a similar number of points. Therefore, we conclude that the simulated galaxy distributions in both models are very filamentary.

The study of percolation through empty cells is a test on the cellular structure. Our results suggest that neither mass nor galaxy distributions in CDM or C + HDM models look like large isolated voids separated by *thin* walls. However, this does not exclude the existence of thin pancakes which do not make a cellular structure or a system of isolated voids separated by *thick* walls, of the kind briefly described above. In this case the density threshold is quite low.

Estimating the percolation threshold in the CfA slices de Lapparent et al. (1991) used a parameter $f_{\rm perc}$ which equals μ_∞ in our notation and $f_{\rm occ}$ corresponding p. However, we wish to warn against a direct comparison of the results, because they used a nonstandard scheme when linking cells.

Our analysis of the percolation properties of a few currently popular cosmological models has shown the potential merits and disadvantages of the percolation method. It also shows additional features discriminating the models. However, to make a practical prediction of the percolation properties of

different cosmological models which can be compared with observations one needs to make additional important steps. More realistic catalogs should be simulated in larger volumes including modeling a real selection function and a redshift space. We are going to do this in a separate paper.

We wish to acknowledge S. Babenko for help in developing the percolation code. We are grateful for the research support from NSF grant AST 90-21414 and NASA grant NAGW-2923 at the University of Kansas. Our computer simulations were done at the National Center for Supercomputing Applications.

REFERENCES

Bhavsar, S. P., & Barrow, J. D. 1983, MNRAS, 205, 61P
Dekel, A., & West, M. J. 1985, ApJ, 288, 411
de Lapparent, V., Geller, M. J., & Huchra, J. P. 1989, ApJ, 343, 1
———. 1991, ApJ, 369, 273
Dominik, K. G., & Shandarin, S. F. 1992, ApJ, 393, 450
Einasto, J., Klypin, A. A., Saar, E., & Shandarin, S. F. 1984, MNRAS, 206, 529
Hockney, R. W., & Eastwood, J. W. 1981, Numerical Simulations Using Particles (New York: McGraw-Hill)
Holtzman, J. A. 1989, ApJS, 71, 1
Kates, R. E., Kotok, E. V., & Klypin, A. A. 1991, A&A, 243, 295
Klypin, A. A. 1987, Soviet Astron., 31, 8
Klypin, A. A., Holtzman, J., Primack, J., & Regos, E. 1992, ApJ, submitted
Mo, H. J., & Börner, G. 1990, A&A, 238, 3
Shandarin, S. F. 1983, Soviet Astron. Lett., 9, 104
Shandarin, S. F., & Zel'dovich, Ya. B. 1989, Rev. Mod. Phys., 61, 185
Stanley, H. E. 1971, Introduction to Phase Transitions and Critical Phenomena (Oxford: Clarendon)
Stauffer, D. 1979, Phys. Rep., 54, 1
———. 1985, Introducton to Percolation Theory (London: Taylor & Francis)
Tomita, H., & Murakami, C. 1988, in Dynamics of Ordering Processes in Condensed Matter, ed. S. Komura & H. Furukawa (New York: Plenum), 167
Wentzel, D. G., & Seidel, P. E. 1992, ApJ, 390, 280
Zel'dovich, Ya. B. 1982, Soviet Astron. Lett., 8, 102
Ziman, J. M. 1979, Models of Disorder (Cambridge: Cambridge Univ. Press)