

## **INFORMATION TO USERS**

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**UMI**

A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA  
313/761-4700 800/521-0600



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# Percolation Analysis as a Tool to Describe the Topology of the Large Scale Structure of the Universe

by

Capp Yess

B.S., Bemidji State University, 1980

M.S., Montana State University, 1987

Submitted to the Department of Physics and Astronomy  
and the Faculty of the Graduate School of the University  
of Kansas in partial fulfillment of the requirements for  
the degree of Doctor of Philosophy

Sergei F. Shandarin

Sergei F. Shandarin, Professor of Physics & Faculty Advisor

Adrian L. Melott

Adrian L. Melott, Professor of Physics

Diss

1997

Y477

(High res)

Hume A. Feldman

Hume A. Feldman, Professor of Physics

Thomas E. Cravens

Thomas E. Cravens, Professor of Physics

Marlin D. Harmony

Marlin D. Harmony, Professor of Chemistry

16 April '97

Date Defended

R00309 38568

MAY 18 1997

**UMI Number: 9811341**

---

**UMI Microform 9811341  
Copyright 1997, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized  
copying under Title 17, United States Code.**

---

**UMI**  
300 North Zeeb Road  
Ann Arbor, MI 48103

# Abstract

Capp Yess, April 1997

University of Kansas

Percolation analysis is the study of the properties of clusters. In cosmology, it is the statistics of the size and number of clusters. This thesis presents a refinement of percolation analysis and its application to astronomical data. An overview of the standard model of the universe and the development of large scale structure is presented in order to place the study in historical and scientific context. Then using percolation statistics we, for the first time, demonstrate the *universal* character of a *network* pattern in the real space, mass distributions resulting from nonlinear gravitational instability of initial Gaussian fluctuations. We also find that the maximum of the number of clusters statistic in the evolved, nonlinear distributions is determined by the effective slope of the power spectrum.

Next, we present percolation analyses of Wiener Reconstructions of the IRAS 1.2 Jy Redshift Survey. There are ten reconstructions of galaxy density fields in real space spanning the range  $\beta = 0.1$  to  $1.0$ , where  $\beta = \Omega^{0.6}/b$ .  $\Omega$  is the present dimensionless density and  $b$  is the linear bias factor. Our method uses the growth of the largest cluster statistic to characterize the topology of a density field, where Gaussian randomized versions of the reconstructions are used as standards for analysis. For the reconstruction volume of radius,  $R \approx 100h^{-1}$  Mpc, percolation analysis reveals a slight ‘meatball’ topology for the real space, galaxy distribution of the IRAS survey.

Finally, we employ a percolation technique developed for pointwise distributions to analyze two-dimensional projections of the three northern and three southern slices in the Las Campanas Redshift Survey and then give consideration to further study of the methodology, errors and application of percolation. We track

the growth of the largest cluster as a topological indicator to a depth of  $400 \text{ h}^{-1}$  Mpc, and report an unambiguous signal, with high signal-to-noise ratio, indicating a network topology which in two dimensions is indicative of a filamentary distribution. It is hoped that one day percolation analysis can characterize the structure of the universe to a degree that will aid theorists in confidently describing the nature of our world.

This work is dedicated to:

Tanner, who will hopefully benefit from it, and  
Manju, who paid for it.

# Acknowledgments

I would like to first thank my advisor, Sergei Shandarin, who taught me more than these pages could ever convey and Adrian Melott, who, in his own way, counseled and encouraged me. I thank my friend Randy Splinter for all his help, and Jenny Pauls for her patience. I thank my mother, father and siblings for understanding and not ridiculing me every time I went back to school. Lastly, I need to thank my first mentor, Ron Anderson, who showed me what a fine teacher is, and Lee Lindblom, who always gave me thoughtful advise which I very seldom followed.

# Contents

<b>1 Overview of the Standard Model of Cosmology and Development of Large Scale Structure</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 The Big Bang and the Early Universe . . . . .	5
1.3 Development of Structure . . . . .	7
1.4 Percolation . . . . .	11
<b>2 Universality of the Network and Bubble Topology in Cosmological Gravitational Simulations</b>	<b>14</b>
2.1 Introduction . . . . .	14
2.2 Method . . . . .	20
2.2.1 Filling Factor . . . . .	20
2.2.2 Largest Structures . . . . .	22
2.2.3 Number of Clusters and Voids . . . . .	23
2.2.4 Mixing Phases . . . . .	24
2.3 Percolation in Grid Gaussian Fields . . . . .	25
2.4 N-Body Models . . . . .	29
2.4.1 Normalization . . . . .	30
2.5 Nonlinear Gravitational Distributions . . . . .	31
2.5.1 Largest Cluster and Largest Void . . . . .	31
2.5.2 Density Contrast of the Largest Cluster . . . . .	33

2.5.3	Number of Clusters and Voids . . . . .	35
2.6	Discussion . . . . .	37
<b>3</b>	<b>Percolation Analysis of a Wiener Reconstruction of the IRAS 1.2 Jy Redshift Catalog</b>	<b>41</b>
3.1	Introduction . . . . .	41
3.2	The Density Fields . . . . .	44
3.2.1	Wiener Reconstructions . . . . .	44
3.2.2	Gaussian Randomizations . . . . .	47
3.3	Percolation Method . . . . .	47
3.4	Results . . . . .	49
3.4.1	Largest Structures . . . . .	50
3.4.2	Number of Clusters . . . . .	56
3.5	Conclusions . . . . .	56
<b>4</b>	<b>Detection of Structure in the Las Campanas Redshift Survey</b>	<b>59</b>
4.1	Introduction . . . . .	59
4.2	The LCRS and Poisson Standards . . . . .	60
4.3	Percolation . . . . .	63
4.4	Results and Conclusions . . . . .	65
<b>5</b>	<b>Conclusion</b>	<b>70</b>
5.1	Summary of Results . . . . .	70
5.2	Error Analysis and Percolation . . . . .	72
5.3	Future Applications . . . . .	72

# List of Figures

1.1	The figure above was produced by the COBE science team and shows three false color images of the sky as seen at microwave frequencies. The orientation of the map is such that the plane of the Milky Way runs horizontally across the center of the image. The map shows the microwave sky after the any dipole anisotropy has been subtracted. This removal eliminates the major fluctuations in the map: the ones that remain are thirty times smaller. On this map, the hot regions (light) are 0.0002 Kelvin hotter than the cold (dark) regions. . . . .	3
1.2	A two-dimensional representation of the Las Campanas Redshift Survey. Note the apparent filaments and voids. . . . .	5
2.1	The relationship of the cumulative distribution function (filling factor) to the density contrast. Solid lines are distributions derived from N-Body simulations, and dotted lines are the Gaussian randomizations. One sigma error bars are contained within the width of the lines. . . . .	21

2.2 Percolation in Grid-Gaussian fields with $k_{cutoff} = 16$ and $32$ (where $k_c = 32$ is the Nyquist cutoff). The top panels show the volume of the largest structure in filling factor units. Short dashes display $1\sigma$ dispersion for all graphs. The number of clusters statistic (normalized to the cutoff volume) is shown in the middle panels. Maximum values are marked with dots. The bottom panels display statistics of the density peaks: theoretically calculated cumulative density of peaks (left), and the number of peaks to cluster ratio for a given filling factor (right). . . . .	27
2.3 Percolation of nonlinear density fields derived from N-body simulations and the Gaussian randomizations. Cluster (over-dense) percolation is shown by solid lines, void (under-dense) percolation by dashed lines, and dots represent the percolation of the randomized fields. The volumes of the largest structures are plotted in units of the filling factor. Short dashes reflect dispersion at the $1\sigma$ level. . . . .	31
2.4 The mean densities of the largest cluster produced in N-body simulations compared to the normalized volumes. Where visible small dashes represent $1\sigma$ dispersions. The thin lines are the value of the density contrast supplied for comparison. . . . .	34
2.5 The number of clusters per Nyquist volume ( $\lambda_{Ny}^3$ ) for a given filling factor. Solid lines are statistics for clusters, dashed lines for voids, and dots for Gaussian randomizations. Short dashes represent $1\sigma$ level dispersions. . . . .	36

3.1 Largest Structure Statistic for IRAS reconstructions over the range $0.1 \leq \beta \leq 1.0$ , where $\beta = \Omega^{0.6}/b$ . Lines are in triplets, except $\beta = 1.0$ , with lighter lines representing larger $\beta$ : solid ( $\beta = 0.1, 0.2, 0.3$ ), dots ( $0.4, 0.5, 0.6$ ), dash ( $0.7, 0.8, 0.9$ ) and dot-dash ( $1.0$ ). The top panels correspond to the measurements in a cube of length $200 h^{-1} \text{ Mpc} = 64$ mesh units on a side, and the bottom panels to measurements in a sphere $30 \text{ mu}$ in radius. . . . .	51
3.2 Largest Structure Statistic for N-body simulations described initially by a spectral index $n = -1$ under a variety of conditions. The top panels report results from the analysis of density fields derived from the initial N-body simulations; whereas, the middle and lower panels display the results from analysis of reconstructions of the density fields. Solid lines indicate cluster (over-dense) percolation while dashed lines indicate void (under-dense) percolation. Light lines represent Gaussian randomizations. Error bars are $1\sigma$ deviations over four realizations in all cases, although commonly they are smaller than the line thickness. . . . .	54
3.3 Largest Structure Statistic for selected IRAS reconstructions including Gaussian randomizations. Solid lines indicate cluster percolation while dashed lines indicate void percolation. Light lines represent Gaussian randomizations. . . . .	55
3.4 Number of Clusters Statistic for Wiener Reconstructions of N-body simulations ( $n = 0, -1, -2$ ) and IRAS data ( $\beta = 0.1, 0.5, 1.0$ ). Solid lines represent clusters and dashed lines represent voids, where light lines indicate Gaussian randomizations. Error bars for the $n = -1$ case represent $1\sigma$ deviations over four realizations. . . .	57

4.1	A representation of the Las Campanas Redshift Survey Slices and the two dimensional sectors onto which the galaxy positions are projected. . . . .	61
4.2	Map of the Las Campanas Redshift Survey Slice. North 1. There are 4495 galaxies plotted (initial Filling Factor of 0.016). To the side is the corresponding Poisson Distribution and below the distributions corrected for projection and selection and projection effects sequentially. The axis units are Mpc. . . . .	62
4.3	The Largest Cluster Statistic (LCS) results for the LCRS slices. The top panels are the results for the full slices ( $R \leq 600$ Mpc), the middle and bottom panels are the results for the magnitude and volume limited samples respectively ( $200 \leq R \leq 400 h^{-1}$ Mpc). The solid lines are the survey results (1 lightest, 3 darkest, for both north and south slices). The dashed lines are the LCS for corresponding Poisson Distributions corrected for selection and projection effects while the dotted lines are uncorrected Poisson Distribution results. The error bars are one sigma deviations over four realizations in all cases. . . . .	66
4.4	Maps of the volume limited distributions for all Las Campanas Redshift Survey slices ( $200$ Mpc $\leq R \leq 400 h^{-1}$ Mpc). The axis units are Mpc. . . . .	68

# List of Tables

2.1	Maxima of the Number of Clusters Statistic (Estimated Slopes) . . . . .	37
2.2	Volume Fractions in Clusters/Voids/Gauss (except the largest structure) . . . . .	40
4.1	LCRS Slice Data . . . . .	65

# **Chapter 1**

## **Overview of the Standard Model of Cosmology and Development of Large Scale Structure**

### **1.1 Introduction**

The desire, the necessity, for humans to contemplate their surroundings for religious, social and scientific purposes is universal and timeless. All people have their myths, religion, social constructs and sense of aesthetics interwoven with their science resulting in a unique cultural identity and perspective of the world. At present, the understanding of the organization and development of the universe has come to be governed by principles which emphasize empirical observations, internal consistency of reasoning, comprehensiveness and explanatory power. This thesis is a continuation of that quest for a more refined, more encompassing idea of our natural world. In particular, this thesis presents percolation analysis; its evolution, methodology and application as a tool to describe the distribution of galaxies (mass) in our local universe. As described by Stauffer & Aharony (1992) percolation theory deals with the statistics of groups of neighboring occupied sites

in a distribution. In cosmological applications occupied sites represent the positions of galaxies in an astronomical survey.

Modern cosmology was born in 1965 with the discovery and explanation of the cosmic microwave background radiation (CMBR). The remnant of the light energy born of the big bang and predicted in the late 1940's by George Gamow and collaborators, established that the universe was expanding and put an end to the competing theory of a steady state universe. The notion of an expanding universe, so foreign to Einstein that he modified the general theory of gravitation to negate the idea, already had its empirical roots in the work of Edwin Hubble. Hubble demonstrated in the 1920's that galaxies were receding from us at a rate that is proportional to their distance. The ratio of the velocity of recession of an object to its distance is called the Hubble constant. This relationship allows for an object's distance from us to be determined if its recession velocity is known. Recession velocities are determined by the amount the light from the object has been redshifted. Another empirical feature of the universe is its degree of isotropy and homogeneity. The degree of isotropy of the universe is measured directly by the difference in temperature of the CMBR over different regions of the sky. Anisotropies in the x-ray background of distant radio sources are also an indication of this feature. The temperature anisotropies of the CMBR also give an upper bound to the inhomogeneities in the density of the universe at the epoch of last scattering. The CMBR along with supporting data from galaxy counts and the study of the peculiar velocities of galaxies reveals a universe that is isotropic and homogeneous to a degree better than one part in ten thousand over many distance scales. However, deviations have been detected (Smoot et al. 1992) and confirmed (Ganga et al. 1993) at the level of  $10^{-5}$  K on an angular scale of  $10^\circ$  (see Figure 1.1).

Other measurable quantities of the universe by direct or indirect methods are its age, density and primordial abundances of light chemical elements. If the

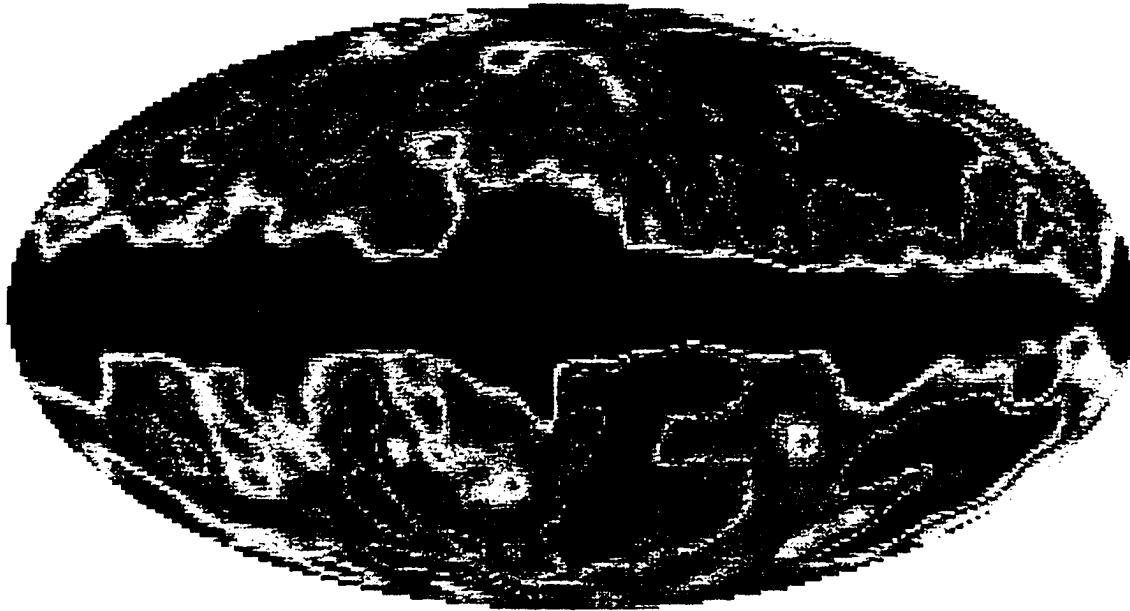


Figure 1.1: The figure above was produced by the COBE science team and shows three false color images of the sky as seen at microwave frequencies. The orientation of the map is such that the plane of the Milky Way runs horizontally across the center of the image. The map shows the microwave sky after the any dipole anisotropy has been subtracted. This removal eliminates the major fluctuations in the map: the ones that remain are thirty times smaller. On this map, the hot regions (light) are 0.0002 Kelvin hotter than the cold (dark) regions.

universe is expanding the inverse of the Hubble constant should give an estimate of its age. This figure (10-15 billion years) is corroborated by dating radioactive elements, determining the ages of the oldest stars, and calculating the cooling time for hot gas in clusters and white dwarf stars. Once a big bang scenario is generally accepted the era of an initial hot, dense ‘plasma’ of elementary particles is unavoidable. Particle physicists have determined the ratios of the elements that would have been formed in the first 100 seconds of the universe. To date there is considerable agreement between the predicted and observed values of the primordial species. The matter predicted by particle physicists is baryonic and by modern calculations should amount to no more than 16% of the mass necessary to halt the cosmic expansion. The average density of the universe is observable by a number of methods covering large and small cosmic scales. Two

important findings come from the research of universal density, baryonic matter can not account for all the mass that is indirectly observed and all matter observed by direct or indirect means amounts to about 20% of the critical density of the universe. One of the consequences of these studies is the prediction of dark matter of various forms. The implications of dark matter for percolation analysis are considerable and one avenue cosmologists have taken to address the problem is to assume that the distribution of dark matter traces the luminous matter to some degree given by a biasing factor. This bias factor relates the statistical properties of the galaxy distribution to the underlying mass distribution. For instance the linear bias factor,  $b$ , can be represented as the ratio of the variance in the galaxy counts to the variance in the mass contained in spheres of radius  $R = 8h^{-1}$  Mpc.

$$b^2 = \sigma_8^2(\text{galaxies})/\sigma_8^2(\text{mass}). \quad (1.1)$$

and is at present undetermined but generally assumed to be in the range  $0.0 \leq b \leq 2.0$ .

The parameters outlined above reflect early or initial conditions of the universe. Observations also need to be interpreted in order to describe the present state of the universe. Statistical methods such as correlation functions and percolation analysis are one way of interpreting galaxy distributions or density fields produced from galaxy surveys. Through examination of the types and positions of galaxies, clusters and superclusters of galaxies can be defined and located (see Figure 1.2). Along with voids and dark matter, these groupings of galaxies constitute the large scale structure of the universe. The study of the large scale structure of the universe is approached on three fronts: (i) theories of the growth of large scale structure due to gravity and other physical influences, (ii) details of the initial density perturbations that act as the initial conditions for the theories and as seeds of structure formation, (iii) quantitative descriptions of the present mass distribution in the universe. Percolation analysis has the potential

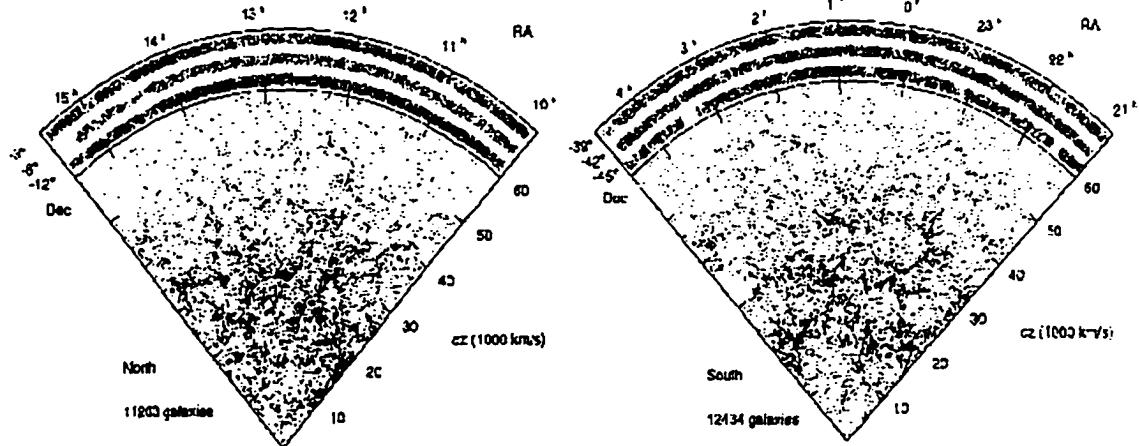


Figure 1.2: A two-dimensional representation of the Las Campanas Redshift Survey. Note the apparent filaments and voids.

to quantitatively describe the detectable mass distribution of our local universe and characterize the initial density fluctuations.

## 1.2 The Big Bang and the Early Universe

By ignoring the small scale deviations from uniformity (stars and planets), we can treat the mass of the universe as a gas of particles (galaxies). By invoking a high degree of symmetry in the form of isotropy and homogeneity, the geometry of spacetime can be greatly simplified. Another common simplification of the Einstein's equations is the use of comoving coordinates <sup>1</sup>. Using comoving coordinates allows for the galaxies to have fixed intervals between them and the expansion of the universe results from a change in the metric of spacetime. If all clocks in the universe are synchronized such that they read zero at the instant of the big bang the distance interval that follows from the curvature tensor that is compatible with the above stipulations is:

---

<sup>1</sup>Comoving coordinates are also a common feature in the set-up conditions of N-body simulations.

$$dl^2 = dr^2/(1 - Kr^2) + r^2d\theta^2 + r^2\sin^2\theta d\phi^2 \quad (1.2)$$

Where  $K$  ( $K = 1/a^2$ ) is the curvature of the three sphere described by this geometry, and  $a$  is the time dependent scale factor. In this development,  $K$  has three value ranges that separate the three classes of curvature and result in distinctly different geometries. For positive  $K$  the universe is finite and closed, for negative  $K$  the universe is infinite and open, and for  $K$  equals zero the universe is infinite and open. To write down and solve Einstein's equations for the universe the stress-energy tensor must be specified. This tensor is a function of the mass density of the universe, the radiation pressure and possibly a non-zero cosmological constant. Various models of the universe are constructed by specifying values of the stress-energy terms. Einstein's equations can then be solved and cast in forms that emphasize the observable cosmological parameters: the mass density, the Hubble constant, the deceleration parameter and the age of the universe.

Besides the expansion of the universe, another consequence of the big bang theory is that the universe is cooling and therefore was hotter at earlier times. The cooling of the constituents of the universe plays an integral role in the character of the universe today. In the make-up of present day matter and the very forces of nature, temperature has been the critical parameter. In the early history ( $t \leq 10^5$  years), matter and radiation were in thermal equilibrium and the universe consisted of a homogeneous plasma. This era ended when the universe was cool enough for ionized hydrogen to change into neutral hydrogen. Today, we live in a matter dominated universe where matter and radiation have been decoupled since the recombination of hydrogen. The free radiation from that time is witnessed today in the form of the microwave background radiation. Before decoupling, however, the universe went through many epochs associated with the different forms of matter that existed at the prevailing temperature.

Particle physicists have detailed a consistent history of the universe starting

at  $t = 10^{-43}$  seconds to the present. Before this time we believe the universe was dominated by quantum fluctuations in the geometry, about which we know very little. At  $t = 10^{-43}$  seconds the Hadron Era began. At the temperatures that define this era only the most basic of particles and their anti particles exist. The constituents cool until quarks combine to form a small excess of protons, neutrons and pions. With the annihilation of most of the baryons, the universe enters the Lepton Era ( $t = 10^{-4}$  seconds). The universe consists in this phase until the leptons annihilate to leave a small excess of electrons, muons, taus and neutrinos. By the time the universe is ten seconds old most of the matter in the universe has annihilated with its associated antimatter and the universe has entered the electrically neutral Radiation Era. The universe stays radiation dominated for roughly  $10^{12}$  seconds and cools from  $10^9$  to  $10^4$  K while the nuclei of the lightest primordial elements form. At approximately  $t = 10^{12}$  seconds the energy density of radiation dropped below the energy density of matter and the Matter Dominated Era began. Matter and radiation were still in thermal equilibrium, however, until recombination when each was free to cool separately. As long as matter and radiation were coupled, any small irregularities in the matter density were held in check by the diffusion of photons. But since decoupling, perturbations in the matter density have been free to grow by gravitational attraction into the structure of the present universe.

### 1.3 Development of Structure

An important problem in the theory of structure formation is the character of the initial perturbations in mass density that give rise to the complexity we see today. How could initial fluctuations in density be present in such a hot and homogeneous gas of particles? There are two prominent scenarios for the production of fluctuations in the density field. One scenario relies on the phase transitions that

assuredly accompanied the spontaneous breakdown of symmetries that produced the independent forces of nature. One consequence of phase transitions would have been cosmic strings, topological defects with considerable effective mass and extent. As these strings moved through the universe they would create wakes which would perturb the mass density and initiate fluctuations.

The other scenario is an inflationary epoch in the history of the early universe. This solution has many appealing features and alleviates many of the shortcomings of the standard model. The basic idea of inflation is that there was an epoch in the early universe when the vacuum energy associated with a scalar field was the dominant form of energy. This energy can be represented by the lambda term introduced by Einstein in his equations. Under the conditions of inflation, a small, smooth, and causally coherent region of the universe could grow exponentially to encompass the entire universe today. This process supplies a rational as to why the universe is homogeneous and isotropic, why topological defects, such as magnetic monopoles, have not been observed, and how initial small scale inhomogeneities in the matter density could have existed at the decoupling era. Specifically density perturbations would be the result of quantum fluctuations in the scalar field. These small scale fluctuations would be expanded to cosmological proportions by the end of the inflationary era.

Because there is no convincing theoretical basis for any one particular theory of perturbation generation, it is usually assumed that the power of primordial fluctuations is best represented by a power law,  $P(k) \propto k^n$ . In theory, observations of the large scale structure can lead to a determination of the spectrum of initial density perturbations. This requires a good understanding of structure formation and good observational data. Percolation analysis is sensitive to the slope of the power spectrum if the initial perturbations have a random Gaussian phase distribution. The fluctuations in density that arise from the general inflationary scenario are thought to be random in their phase distribution. Therefore,

percolation analysis has the potential to estimate the spectral index of the power spectrum of primordial perturbations.

At this point we do not know the total amount of matter in the universe or its composition, nor do we know the spectrum of primeval perturbations. but we do have a record of the perturbations that existed at the time of recombination. The smoothness of the CMBR restricts the amount of structure formation up to that time, but does allow for the seeds of structure. The current picture of structure development from small inhomogeneities in the mass density is based on the notion of gravitational instabilities. In a medium of uniform density any local density excess will attract matter by its own gravitation and this effect will accelerate. However, these density enhancements are opposed by the pressure of the gas, so it takes a minimum mass to overcome the stabilizing pressure forces. This minimum mass is called the Jeans mass, and together with the speed of sound in the gas and the average density an approximate radius of the collapsing region can be computed.

$$\lambda_J = (\sqrt{24}/5)v_s(\pi/G\rho)^{1/2} \quad (1.3)$$

For a universe with random inhomogeneities the consequence of this theory is a distribution of discrete objects with sizes greater than or equal to the Jeans mass. There are two possible types of fluctuations that can exist within the framework of general relativity, curvature and isocurvature. Curvature, sometimes called adiabatic, fluctuations are characterized by fluctuations in the local energy density and constant entropy; isocurvature (isothermal) fluctuations are characterized by fluctuations in the local entropy with constant energy density and temperature.

Unfortunately, if baryonic matter is the only form of matter in the universe. the time between matter-radiation equality and recombination requires inhomogeneities larger than those measured in the CMBR to produce the overdensities associated with galaxies and clusters today. One possible solution to this problem

is that nonbaryonic, dark matter decoupled from radiation earlier than baryons did and the fluctuations in dark matter accelerated the large scale structure development. Depending upon whether the particles comprising the dark matter are relativistic or not when they decouple from the other constituents of the universe, they are called hot or cold dark matter. The two types of dark matter and the two types of fluctuations defined above provide very different pictures of the evolution of perturbations. Adiabatic fluctuations and hot dark matter suppress the formation of overdensities on small scales which leads to large pancakes of gas being the first objects to form in the universe. In the adiabatic case, photon diffusion disperses the matter in overdense regions and the free streaming of neutrinos in the hot dark matter scenario erases small scale inhomogeneities. The consequence of these influences is that structures on the scale of clusters and superclusters of galaxies form first in the universe. These structures break apart into galaxies in what is called the top-down scenario. In contrast, there are no analogous damping mechanisms in the cases of cold dark matter and isothermal fluctuations, so in these pictures small scale structures form first (globular clusters and dwarf galaxies). In these bottom-up scenarios, larger structures coalesce thru the gravitational attractions of nearby galaxies. In addition, the bottom-up scenarios are associated with initial power spectra with positive spectral index where there is relatively more power on small scales. Whereas, the top-down picture follows from an initial power spectra with negative spectral index indicating relatively more power on large scales.

An important feature of the different classes of hierarchical development is that different degrees of connectedness result in the distribution of mass and galaxies<sup>2</sup> that form under the different scenarios. The differences in the distributions are topological distinctions. There is a broad range of characterization for the topol-

---

<sup>2</sup>A distinction is made here between mass and galaxy distributions to allow for the possibility that galaxies are only tracers of a more extensive underlying mass distribution. In this picture galaxies form only at the very highest overdensity regions.

ogy of mass and galaxy distributions stretching from filamentary to clumpy. A filamentary topology would describe a mass network consisting of connected structures. In three dimensions, this would be consistent with connected pancake and cigar shaped structures. A clumpy topology, sometimes referred to as a meatball topology, describes matter that has coagulated into isolated regions surrounded by voids. Percolation analysis is a measure of the connectedness of a distribution and is therefore sensitive to its topology. Percolation can then be used as a quantitative tool to describe and distinguish a distribution as well as estimate the slope of its power spectrum. In this way percolation analysis helps to solve two of modern cosmology's problems, the nature of the primordial fluctuations and the development of the large scale structure of the universe.

## 1.4 Percolation

Percolation was originally developed to study the connections of molecules in the creation of polymers. Scientists in the 1940's were interested in the process of *gelation*, that is the formation of chemical bonds that span an entire system. Since that time percolation has been adapted for and applied to many scientific and engineering topics. Each application of percolation has its own intents and dynamical parameters, but all rely on the statistical analysis of the properties of clusters to produce results. The parameters that are commonly used in cosmological percolation are (1) the percolation variable, which is a property of the distribution that defines the cluster boundaries; (2) the filling factor (FF), the fraction of survey volume contained in the clusters for a given value of the percolation variable; (3) the largest cluster statistic (LCS), the relative volume of the largest cluster to all other clusters (see Klypin & Shandarin 1993); and for percolation of a mass or density field (4) the number of clusters statistic (NCS), the normalized number of clusters at a given value of the percolation variable.

The general idea of percolation is that as some variable of a system changes, the clusters of the system undergo a transition from being isolated from each other to being joined with each other until they span or fill their space. The general percolation method for pointwise distributions follows these steps: (1) superimposing a grid on the survey region, (2) labeling grid cells that contain galaxies as filled, (3) identifying clusters by a nearest neighbor scheme, (4) determining the initial values of the percolation parameters (LCS and FF), (5) centering circles at each galaxy location, (6) incrementally increasing the radius of the circles and relabeling enveloped grid cells as filled while tracing the growth of clusters by the nearest neighbor definition, (7) at each increment of the circle radius, recalculating the filling factor and the largest cluster statistic through the percolation transition. Percolation of density or mass fields have a similar and analogous methodology. In the 1980's cosmologists began adapting the analysis to study the mass and galaxy distributions derived from N-body simulations and astronomical surveys (for example, Einasto et al. 1983). In the cosmological case the system being analyzed is the distribution of galaxies, the parameter being varied is the size or extent of the galaxies, and the transition from isolated groups of galaxies to a network of galaxies is a topological transition.

The main objective of this effort has been to characterize the topology of the distribution of galaxies in the observable universe as being connected or clumpy. The determination of the topology has broad implications for theories of large scale structure development. In this context, percolation analysis has the potential to determine whether the top-down or bottom-up scenario of structure development is more likely and suggest the nature of the power spectrum of initial fluctuations. The contents of this thesis present the state to which that end has been achieved to date.

Chapter 2 continues a refinement of the percolation parameters and methods of characterization of distributions that were started by cosmologists more than

a decade ago. The discovery that the topology of mass distributions derived from N-body simulations with initial conditions described by a broad and reasonable range of power spectra result in a network topology is presented. In addition, the sensitivity of percolation analysis to the spectral index of the power spectrum is explained and demonstrated. In chapters 3 and 4 percolation analysis is applied to astronomical data. A Wiener reconstruction of the IRAS 1.2 Jy redshift catalog is analyzed as an example of a mass distribution. The results and difficulties of this study are enumerated in chapter 3. Chapter 4 gives the results of the pointwise percolation of the Las Campanas Redshift Survey, a state of the art galaxy survey in terms of coverage volume and sample size. A summary of the conclusions and results of this thesis are presented in chapter 5 along with an assessment of future developments.

# **Chapter 2**

## **Universality of the Network and Bubble Topology in Cosmological Gravitational Simulations**

### **2.1 Introduction**

The topology of the galaxy distribution can provide important clues to the formation of large-scale structure in the universe and the nature of the initial density fluctuations (Shandarin & Zel'dovich 1983). The recent compilation of large galaxy surveys, large in galaxy numbers and survey volumes, has prompted a new round of topological studies of the structure of the universe. The amount of information in studies like the IRAS and CfA Redshift Surveys, and the upcoming Sloan Digital Survey suggests that the major problems of discreteness, boundary effects, and local inhomogeneity that have plagued topological analysis to date may be overcome.

Two types of methods have been developed and employed to quantitatively

assess the topology of the galaxy distributions of these surveys. The first one is based on the evaluation of the mean Euler characteristic,  $\chi(\nu)$ , or the genus curve,  $g(\nu)$ , as functions of the density contrast:  $\nu = \delta/\sigma_8$  ( $g = -\chi/2$ ) (Doroshkevich 1970; Bardeen et al. 1986; Gott, Melott & Dickinson 1986)<sup>1</sup>. The genus curve method has been utilized for studies of N-body simulations (Gott, Weinberg, & Melott 1987), nonlinear gravitational evolution (Melott, Weinberg & Gott 1988) and many galaxy catalogs (see e.g., Weinberg, Gott & Melott 1987; Gott et al. 1989; Moore et al. 1992; Vogeley et al. 1994).

The other method is based on percolation theory. Percolation is the study of the number and various properties of "clusters".<sup>2</sup> In 1982, Zel'dovich noticed that the percolation properties of the nonlinear density distribution in the HDM (Hot Dark Matter) model are very different from that in the initial Gaussian field. In particular, the formation of a percolating cluster (the cluster spanning through the entire system) happens more effectively than that in Gaussian fields. He also suggested characterizing the topology of nonlinear density distributions by their percolation thresholds (Zel'dovich 1982). Following Zel'dovich's idea, one of the authors of this paper (S.Sh.) suggested using percolation properties of the *galaxy distribution* as an objective quantitative measure of the topology of the large-scale structure and also as a discriminator between cosmological models (Shandarin 1983; Shandarin & Zel'dovich 1983).

A percolation technique was utilized in the study of the CfA I catalog (Einasto et al. 1984) and showed that the large-scale distribution of galaxies had a network structure. Theoretical studies of models with a power law initial spectra showed that the  $n = -1$  model clearly percolated better than the  $n = 0$  model, and in the  $\Omega = 1$  universe, the  $n = -1$  model was in agreement with observations (Gott

---

<sup>1</sup>For a general review of this method see e.g. (Melott 1990)

<sup>2</sup>In the absence of a better term, we label as "clusters" the high density regions bounded by the surfaces of chosen constant density and "voids" as bounded regions of low density. These terms are not to be confused with the astronomical terms, voids and clusters of galaxies.

& Rees 1975; Bhavsar & Barrow 1983). The percolation method also showed that the CDM (Cold Dark Matter) model appeared to have a connected rather than clumpy density distribution (Melott et al. 1983; Davis et al. 1985). Later it was pointed out that the major disadvantage of any percolation technique was the dependence of the percolation thresholds on the mean density of the sample (Dekel & West 1985) which made it difficult to calibrate for sparse samples. Similarly, we note that at present some believe that sparse samples can only be reliably used for the estimation of the two-point correlation function (Bouchet 1995). We elaborate on this question below.

For continuous fields such as density fields the independent parameter is the density threshold. A cubic lattice is superimposed on the field, and lattice cells with densities above the threshold are tagged as over-dense while cells below the threshold are labeled under-dense. Over-dense cells are considered clusters, and can merge with other over-dense cells to become larger clusters by satisfying a nearest neighbor condition. The nearest neighbor criterion we employ is that over-dense cells must share a common side. On a simple cubic lattice, this means that each cell can have up to six nearest neighbors<sup>3</sup>. As the value of the threshold density is decreased, more cells will be tagged as over-dense, and clusters will grow in number and/or size. This process proceeds until the largest cluster spans the available space and percolation is said to have occurred. Void percolation is accomplished in an analogous fashion except that the density threshold is initially set at a low value, and voids grow as the threshold value is increased. To quantify the study of the clusters (voids) formed by the above scheme, we trace the value of the three most robust parameters: the cumulative distribution function, referred to as the filling factor; the volume of the largest structure (cluster or void) as a fraction of the corresponding filling factor; and the total number of all clusters and

---

<sup>3</sup>For the versions of percolation analysis utilizing other nearest neighbor definitions or lattice structures see for example, Mo & Börner 1990; de Lapparent, Geller & Huchra, 1991.

voids as a function of the filling factor. The rationale for using the filling factor instead of the density contrast as a fundamental parameter is that the filling factor is normalized allowing a direct comparison of the parametric values of Gaussian and nonlinear distributions as well as the topology of over-dense and under-dense phases. Also we wish to isolate the information stored in the dependence of the filling factor on the density threshold from that stored in the other statistics.

The percolation threshold of a Gaussian distribution, as described above, is believed to coincide with a change in the sign of the genus; however, there is no theorem proving this. Intuitively, it is plausible that percolation thresholds approximately coincide with changes of the genus sign. Sathyaprakash, Sahni & Shandarin (1995) showed that in the nonlinear distributions both transitions happen at close but significantly different filling factors.

One advantage of the genus method is the existence of an analytic expression for Gaussian random fields (Doroshkevich 1970; Hamilton, Gott & Weinberg 1986). Tomita (1986) gave a very elegant analytic expression for the mean Euler characteristic for multi-dimensional Gaussian fields, and recently an analytic expression was obtained in the weakly nonlinear regime (Matsubara 1994). However, one should not forget that the mean genus is a statistical measure and therefore an estimate of the dispersion is needed before it becomes meaningful. The dispersion of the genus for finite samples having finite resolution has not been obtained analytically and can be estimated only from numerical simulations. Percolation parameters are also calculated numerically, but if the dispersion can be estimated, the mean can also be estimated with similar accuracy.

It has been claimed that the percolation thresholds are the most sensitive discriminators of varying models (Shandarin 1983). The recent study of the CfA II catalog using the genus method (Vogeley et al. 1994) supports that suggestion. The authors reduced the information of the genus curve to three numbers one of which was the genus peak width,  $W_\nu = \nu_+ - \nu_-$  (where  $\nu_+$  and  $\nu_-$  are the levels

at which the genus changes sign). Figures 12 through 14 in Vogeley et al. (1994) clearly demonstrate that  $W_v$  has the highest discriminating power among three suggested parameters. However, we still believe that the percolation thresholds as well as  $\nu_+$  and  $\nu_-$  should be interpreted separately because they carry independent information about the topology of the structure.

The major improvement of the percolation technique we report in this paper is related to the development of an extremely efficient code for finding all the clusters at a given density threshold (Stauffer & Aharony 1992; Klypin & Shandarin 1993). This enables us to calculate more parameters with finer density thresholds variations than in earlier studies. In this paper, we shall report the results of studying the nonlinear, density distributions in real space obtained from N-body simulations of the scale free cosmological models:  $P(k) \propto k^n$  with  $n = +3, +1, 0, -1,$  and  $-2$  in the  $\Omega = 1$  universe. The volume of the largest structure will be used in this paper to indicate the onset of percolation and characterize the topology of the distribution. Every generic density distribution at a sufficiently high density threshold will look like a system of isolated (non-percolating) clusters. At some lower threshold the connected system spanning through the entire volume will inevitably form. One possible definition of a network distribution would be a distribution for which percolation occurs at a specified density threshold. In this case the choice of the threshold must be physically justified. We do not know how to select this particular density threshold. Therefore we adopt a different definition of a network distribution. Fields that percolate at lower filling factors than Gaussian fields (by definition structureless) will be interpreted as having a "connected" or network structure; while fields that percolate at higher filling factors will be considered "disconnected" or clumpy. Similarly we label a distribution as having a bubble topology if the under-dense region percolates at a higher filling factor than in a Gaussian field.

It is worth stressing that the terminology in this field is often confusing. We

use the labels clumpy, network, and bubble structure to emphasize the degree of connectedness only. In this paper we ignore the geometrical aspect of the problem. For instance, we do not distinguish between a network made of filaments (quasi one-dimensional objects) or pancakes or sheets (quasi two-dimensional objects). In realistic distributions, it is often impossible to rationally assign a label to the shape of a particular density enhancement. Various shape statistics (Vishniac 1986; Babul & Starkman 1992; Luo & Vishniac 1995) assume that the shapes can be approximated by a triaxial ellipsoid. However, the dynamics of the non-linear gravitational evolution suggests that the first collapsed objects (pancakes) have a bowl like shape (Arnold, Shandarin & Zel'dovich 1982; Shandarin et al. 1995) which is very poorly approximated by a triaxial ellipsoid. In particular, the thickness of such a structure would be highly exaggerated if it is approximated by an ellipsoid. This problem is addressed in a paper by Sathyaprakash, Sahni & Shandarin (1995).

In addition to determining the topology of the density distributions, we will demonstrate a method of estimating the slope of the power spectrum characterizing a distribution. The maximum of the number of clusters (voids) statistic is determined by the effective spectral index at the Nyquist frequency (or the smoothing scale). Potentially, this relation can be used for measuring the slope of the spectrum.

In §2 we will explain the parameters we use to characterize the density distribution in space and discuss their application to the power law Gaussian fields. In §3 we describe the percolation of Gaussian fields, and in §4 we describe the N-body simulations forming the basis for our standardizations. We detail the results from studying the growth of the largest structures and elaborate on the number of clusters results in §5. We summarize our findings in §6.

## 2.2 Method

### 2.2.1 Filling Factor

Percolation theory deals with the number and properties of the clusters. The density threshold,  $\delta_c$ , separating over-dense ( $\delta > \delta_c$ ) and under-dense ( $\delta < \delta_c$ ) regions is assumed to be a free parameter. When analyzing discrete distributions (e.g. galaxy distributions), we assume a smoothing procedure creating a continuous density distribution.

As we mentioned, the density threshold is not a convenient parameter if linear (Gaussian) and nonlinear density distributions or over-dense and under-dense phases are to be compared. Instead we utilize the filling factor to parameterize the density threshold. In this case one can easily compare the properties of clusters with those of voids and also linear and nonlinear density distributions. In a two-dimensional illustration it would be similar to comparing different patterns provided that the same amount of ink was used to make each pattern. For reference we provide the relationship between the filling factor and the density contrast, for the models studied (Figure 2.1). The effect of evolution on the relationship is evident in the graphs and underscores the reasons to use the filling factor as a means of comparison. One sigma error bars are comparable with or smaller than the thickness of the lines and are not shown. Solid lines and dashed lines show the filling factor for the nonlinear distributions and Gaussian fields with corresponding spectra (see section 2.4).

We study various characteristics of a field as a function of the filling factor which is the total volume occupied by the regions having a density above (if we study over-dense regions) or below (if we study under-dense regions) a specified threshold. The filling factor coincides with the “volume fraction” used by Gott, Weinberg & Melott (1987) in their development of the genus curve for topological analysis and by Vogeley et al. (1994). However, instead of expressing it in terms

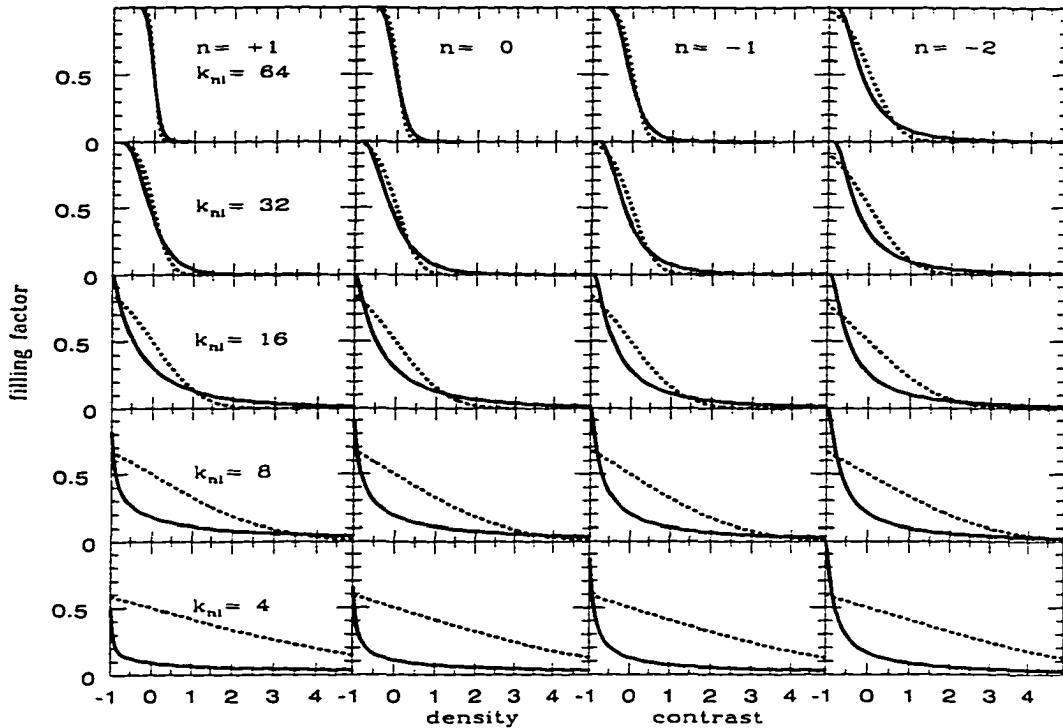


Figure 2.1: The relationship of the cumulative distribution function (filling factor) to the density contrast. Solid lines are distributions derived from N-Body simulations, and dotted lines are the Gaussian randomizations. One sigma error bars are contained within the width of the lines.

of  $\nu$  (the number of standard deviations above or below the mean density of the isodensity contour filling the same volume in a Gaussian field), we use it directly. The filling factor is obviously related to the cumulative distribution function of the corresponding phase:

$$ff(\nu) = \frac{1}{\sqrt{2\pi}} \int_{\nu}^{\infty} e^{-t^2/2} dt. \quad (2.1)$$

Also, in Gaussian fields there is no statistical difference between under-dense and over-dense regions; therefore, the filling factor is the same for both.

## 2.2.2 Largest Structures

Here for definiteness we shall talk about over-dense regions. the under-dense regions can be discussed similarly. For a given density threshold we have a set of regions that differ by volumes, shapes and other parameters. We select the one having the largest volume and call it the largest cluster. It is convenient to measure the largest cluster and void as a fraction of the corresponding filling factor. Thus, if the largest cluster is 0.9 at filling factor of 0.2 it means that the density is higher than the chosen threshold in 20% of the volume and almost all of that volume (90%) is comprised of only one connected region. The absolute volume of the largest cluster is clearly 18% of the total volume of the sample.

In a finite system, (like the density distribution in the N-body simulations or galaxy surveys) if we start from a very high density threshold we do not find any clusters at all. Lowering the threshold we find the first cluster corresponding to the highest density peak. Obviously, it is also the largest cluster. and its volume measured in terms of the filling factor is unity. This is an effect of a finite system and we shall ignore it. In other words one cannot use the statistics at density thresholds which are too high, or filling factors which are too small. At lower thresholds we typically have many clusters. Usually the volume of the largest cluster is negligible compared to the total volume of all clusters (the filling factor). The actual volume of the largest structure (in the units of the filling factor) can be used as a measure of the fairness of the sample: the smaller the largest structure at low values of the filling factor the better the sample. Decreasing the density threshold we reach the situation where clusters begin to merge. As a result the largest cluster becomes a significant fraction of the filling factor. and eventually almost all over-dense regions combine to make just one cluster. At this stage the volume of the largest cluster almost equals the filling factor, and percolation is said to happen. In the past much effort was spent on an accurate measurement of the percolation threshold (Klypin 1987; Dominik & Shandarin 1992; Klypin &

Shandarin 1993). In contrast, we use the *largest structure as a function of the filling factor* as a practical indicator of the percolation transition as suggested by Shandarin (1994, 1995).

The percolation transition is a universal behavior of every non-degenerate system: however, the filling factor at which the transition happens is particular to each system and may vary. The percolation threshold marks a change in the topology of a distribution. Above the percolation threshold ( $\delta_c \geq \delta_p$ ) the over-dense system consists of isolated clusters having finite volumes, and the topology is of a meat-ball or clumpy type (if the under-dense regions are considered it is usually called a bubble topology). Below the percolation threshold ( $\delta_c \leq \delta_p$ ) most of the over-dense volume is in one cluster spanning the entire system, and the topology is labeled a network or sponge topology. Both terms, network topology and sponge topology, correspond to the positive genus, but the term network structure suggests that the percolating structure is “thin” which is a geometrical rather than topological characteristic.

We plot the largest cluster and the largest void volume fractions versus the corresponding filling factor in the same diagram. As we mentioned before, Gaussian fields exhibit no statistical difference between over-dense and under-dense phases and both structures have similar properties. In contrast, for all nonlinear density distributions the largest cluster percolates as easy or easier (that is at smaller filling factors) than the largest void which is a manifestation of non-Gaussianity resulting from nonlinear gravitational clustering.

### 2.2.3 Number of Clusters and Voids

The third type of statistic we use in this paper is the total number of clusters and voids at a given filling factor. The significance of this statistic is its sensitivity to the slope of the spectrum. We shall measure the number of structures per Nyquist volume:  $V_{Ny} = \lambda_{Ny}^3 = 8$  mesh cells. The number of clusters and/or voids typically

grows with the growth of the filling factor until it reaches the value about 0.1-0.13 then it quickly decreases because of merging of the clusters. In the nonlinear distributions the number of clusters is often (but not always) smaller than that of voids at the same filling factors. Again the difference between the number of the clusters and voids is a manifestation of non-Gaussianity of the density field.

#### 2.2.4 Mixing Phases

We produce three distinct types of density fields for analysis in this study: random Gaussian fields with pure power law spectrum, fields derived from N-body simulations with evolved nonlinear power spectrum, and Gaussian fields with nonlinear power spectrum produced by randomizing the phases of the N-body simulation distributions. Percolation of Gaussian fields (with linear and nonlinear power spectrum) produces standards for characterizing the topologies and estimating the spectral indices of density fields in subsequent studies. The analysis of fields generated from N-body simulations checks and calibrates our ability to describe fields by the method outlined above.

In the comparison of simulation fields to Gaussian fields, we wish to avoid grid effects as much as possible. Because grid effects are inherent in the method, our solution is to generate nonlinear Gaussian fields from the simulation fields, thereby, assuring grid effects similar to those displayed in the nonlinear, parent distributions. This also resolves the question of how to calibrate the percolation curves for sparse samples. A good way to do this is to start by Fourier transforming a nonlinear N-body simulation field to  $k$ -space. Then, in the resulting transform generate random phases keeping the amplitude of every wave exactly the same as before, and finally make an inverse Fourier transform resulting in a new Gaussian field having the same power spectrum as the original nonlinear parent field. We ignore the fact that this field may have negative values since we study only the percolation properties parameterized by the filling factor. Since the generated

Gaussian field has the same amplitudes it must be affected by the finite resolution similarly to the nonlinear density fields. This randomized version of the simulation density field is percolated to produce an additional standard for distinguishing the topology of the model and to illustrate the relationship between phase correlations and nonlinearity. The comparison of the nonlinear and randomized (Gaussian) fields provides a measure of the phase correlation in the nonlinear fields. We show and interpret the collective results of the percolation of Gaussian and simulation fields in the following sections.

## 2.3 Percolation in Grid Gaussian Fields

The ability to recognize Gaussianity in a distribution is important for two reasons. First random Gaussian fields are used as benchmarks for describing the topology of density fields derived from survey data, and secondly, it has been demonstrated that random phase density fluctuations could be produced by an inflationary period in the early history of the universe (see e.g. Linde 1990). Importantly, a random Gaussian field is fully distinguished by its associated power spectrum. If gravitational instabilities were the only mechanisms of structure formation and the initial density perturbations were Gaussian, then the initial power spectrum would determine the large scale structure of the universe.

It is worth emphasizing that random fields generated numerically on a grid are only approximately Gaussian. We shall call them grid Gaussian fields. One obvious deficiency of a grid Gaussian field compared to a true Gaussian field is that the probability distribution function is always wrong for sufficiently large values of the field. We will report another less obvious deficiency related to the topology of the field.

The Gaussian fields we employ in this study are created by transforming an array of coefficients, that are random Gaussian distributed, from  $k$ -space to real

space. The range of  $k$  was limited by both the box size (the fundamental mode,  $k_f$ ) and the resolution of the mesh (the Nyquist wavenumber,  $k_{Ny}$ ) such that  $k_f \leq k \leq k_{Ny}$ . We also study fields with cutoffs for which  $P(k) = 0$  for  $k > k_c$ . The results are Gaussian random density fluctuation fields which are homogeneous and isotropic. These properties insure that the associated power spectrum is a function of  $k$  only. The spatial modes of these fields are mutually independent and have random phases. (We exploit this random phase condition of Gaussian fields in our phase mixing of N-body simulations explained above.) It is important to remember that Gaussian fields are by definition structureless.

Figure 2.2 shows several parameters calculated for grid Gaussian fields with power law spectra:  $P(k) \propto k^n$  with the spectral index  $n = -2, -1, 0, +1$ . The fields were generated on a  $64^3$  mesh by Fourier transform of random numbers, so no particles were involved. The results for two cutoff values are shown:  $k_c = 16$ , and  $k_c = 32$  (the Nyquist wavenumber). Four realizations with different random number seeds were generated for each curve, and mean values are plotted. One  $\sigma$  error bars are also shown, but in some cases they are too small to be seen.

The two top panels show the largest cluster measured as a fraction of the filling factor. The left hand side panel shows the grid Gaussian functions with the power law spectra cut off at  $k_c = 16$ , and the right hand side panel shows those without an arbitrary cutoff. According to percolation analysis, percolation in Gaussian fields occurs at  $ff \approx 0.16$  (corresponding to  $\nu = 1$ ) independently of the power spectrum. The onset of percolation is marked by a sharp growth of the largest cluster and corresponds to the change of the topology from meat-ball to sponge type. The left top panel is in agreement with this result, but the right top panel is not. The right panel shows that the percolation threshold depends on the spectral index; the larger the index the higher the percolation threshold. Statistically the difference is highly significant. Thus, the limited resolution of the mesh results in grid Gaussian fields, with no spectral cutoff, having a meat-ball topology shift

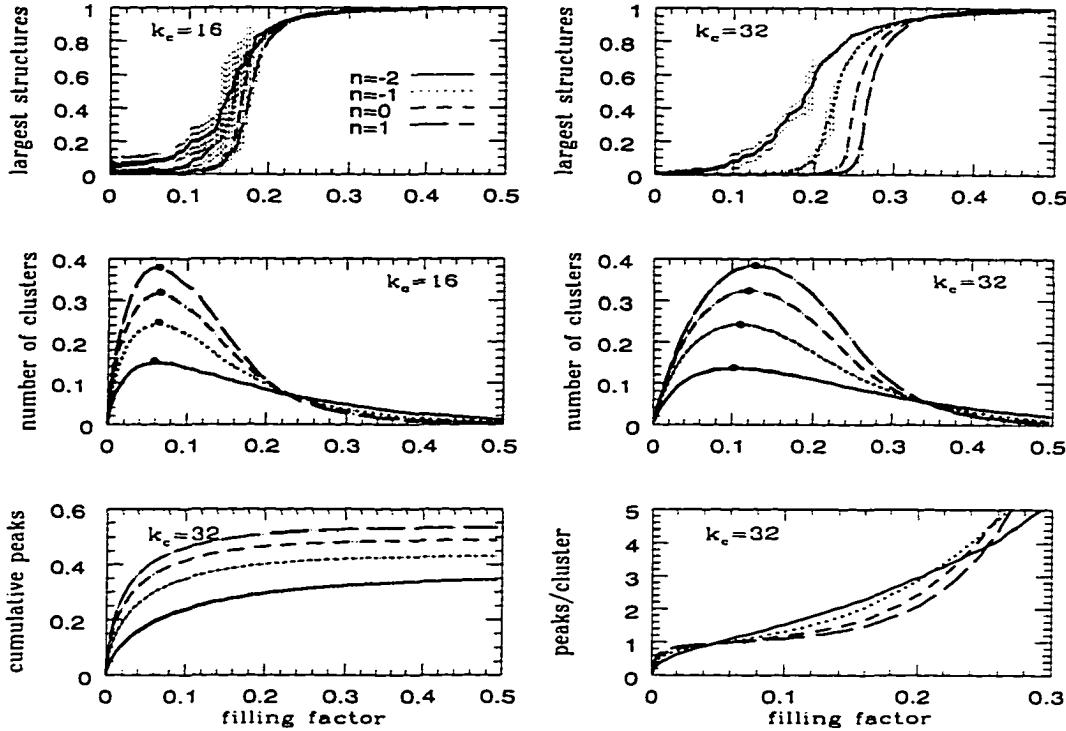


Figure 2.2: Percolation in Grid-Gaussian fields with  $k_{cutoff} = 16$  and  $32$  (where  $k_c = 32$  is the Nyquist cutoff). The top panels show the volume of the largest structure in filling factor units. Short dashes display  $1\sigma$  dispersion for all graphs. The number of clusters statistic (normalized to the cutoff volume) is shown in the middle panels. Maximum values are marked with dots. The bottom panels display statistics of the density peaks: theoretically calculated cumulative density of peaks (left), and the number of peaks to cluster ratio for a given filling factor (right).

indicated by a percolation transition at greater filling factors than true Gaussian fields. This may affect the appearance of the large-scale structure in the N-body simulations especially in those having a small number of particles, making them look more clumpy than they should.

At small filling factors the largest structure must be negligible in sufficiently large samples. The finite size of the largest structure can be used as an internal characteristic of the fairness of a sample. The smaller the largest cluster in the non-percolating regime the better the sample. One can see that the more negative

the spectral index the more difficult it is to satisfy this condition.

The middle panels show the number of clusters per Nyquist volume (8 mesh cells) as a function of the filling factor. The curves reach their maximum at filling factors of about 0.11 (dependent slightly on the spectral index  $n$ ) for the spectra with no arbitrary cutoff ( $k_c = k_{Ny} = 32$ ) and drop with the growth of the filling factor because of the merging of clusters. In the  $k_c = 16$  models the maxima are reached at filling factors ( $ff_{max} = 0.08$ ) which are roughly half the value of the filling factors at the percolation threshold regardless of the spectral index. The value of the maximum (marked by a dot) depends very weakly on the cutoff and is determined only by the spectral index on the scale of the cutoff. This demonstrates that percolation analysis can discriminate between random Gaussian fields typified by different spectral indices (for pure power law initial power spectra) as easily as the maximum of the genus curve.

The left bottom panel shows the cumulative number density of peaks calculated analytically using the equation

$$n_{pk}(ff) = \int_{-\infty}^{\nu} N_{pk}(ff) d(ff) \quad (2.2)$$

(from Bardeen et al. 1986). Since the statistical properties of Gaussian fields are generally reported in terms of the peaks, an interesting quantity is the number of peaks per cluster which can be estimated as the ratio of the number of peaks to the number of clusters. This ratio is shown in the lower right panel. For small filling factors ( $ff \leq 0.02$  or so corresponding to  $\nu \geq 2.5$ ) the number of clusters obtained numerically is unreliable and so is therefore the number of peaks per cluster. In the range  $0.02 \leq ff \leq 0.1$  ( $1.2 \leq \nu \leq 2.5$ ) there is roughly one peak per cluster (however, the exact number depends on the spectral index  $n$ ). Thus we conclude that the numerical result is roughly consistent with the analytical calculation of Bardeen et al. (1986). For larger filling factors the number of clusters drops quickly due to merging into the largest cluster and the ratio of the

number of peaks to the number of cluster grows limitlessly.

## 2.4 N-Body Models

The N-body simulations are produced by a staggered Particle-Mesh code (Melott 1986) with  $128^3$  particles on a  $128^3$  mesh and a corresponding Nyquist wavenumber,  $k_{Ny} = 64$ . The initial conditions are generated by the Zel'dovich approximation (Klypin & Shandarin 1983) such that the initial power spectrum is a simple power law covering the range  $n = 3, 1, 0, -1$  and  $-2$ . The models are allowed to evolve gravitationally until nonlinear effects change the slope of the power spectrum. This change indicates that phase correlations have developed between the originally random initial phases.

The extent of nonlinearity can be characterized by the parameter  $k_{nl}$ , defined by the equation  $\sigma_8^2 = a^2 \int_0^{k_{nl}} P(k) d^3k = 1$ , and in this study we evolve the simulations to values of  $k_{nl} = 64, 32, 16, 8$  and  $4$  (in units of the fundamental frequency). The value of  $k_{nl}$  relates to the scale of structure formation in real space. For a detailed discussion of the N-body simulations see Melott & Shandarin (1993).

Density fields are derived from the above simulations by a cloud in cell method whereby each particle's "weight" is proportionately spread over a  $2^3$  cell volume and rescaled (8:1) to produce a  $64^3$  density field. This method implies some smoothing at small scales but reduces shot noise so that further smoothing is not needed before percolation analysis. An ensemble family of four realizations is produced for each combination of  $n$  and  $k_{nl}$  to give assessments of the one sigma level dispersion for each percolation parameter analyzed.

### 2.4.1 Normalization

It is likely that no single model studied can pretend to explain the real universe. We consider them to be toy models. However, if one wishes to get a rough idea of how they may relate to the real world we provide the following normalizations. We assume that the rms fluctuation in number of galaxies,  $\sigma_g$ , is about unity within spheres of radius  $8 h^{-1} Mpc$ , the rms mass density fluctuation  $\sigma_m$  is parameterized by the “bias factor”,  $b$ , such that  $\sigma_g = b \sigma_m$ . We shall assume that  $b \approx 1$  which is an adequate assumption for these crude estimates. Melott and Shandarin (1993) showed that for the models in question the scale of nonlinearity measured by the top-hat smoothing filter  $R_{TH}$  is approximately two times greater than  $k_{nl}^{-1}$  calculated from the extrapolation of the linear theory (more accurately:  $R_{TH} \approx 1.8 k_{nl}^{-1}$  in the  $n = +1, 0, -1$  models and  $R_{TH} \approx 2.8 k_{nl}^{-1}$  in the  $n = -2$  model). Thus, identifying every stage with the present time one can roughly estimate the size of a mesh cell:  $l_c \approx 25, 12.6, 6.3, 3.1,$  and  $1.6 h^{-1} Mpc$  for  $k_{nl} = 64, 32, 16, 8,$  and  $4$  respectively. In our models the smoothing has been performed with a top-hat filter having a cubic rather than spherical shape which may add an additional factor  $0.6 \approx (4\pi/3)^{-1/3}$  (assuming the volumes of the filters are similar:  $l_c^3 = (4\pi/3)(R_{TH}^{(s)})^3$ ). Therefore, one can view each stage of the evolution of the models as the density distribution seen after smoothing with a top-hat filter of radius  $R_{TH}^{(s)} \approx 16, 8, 4, 2,$  and  $1 h^{-1} Mpc$  within the volumes of  $(64 l_c)^3 \approx (1600 h^{-1} Mpc)^3, (800 h^{-1} Mpc)^3, (400 h^{-1} Mpc)^3, (200 h^{-1} Mpc)^3,$  and  $(100 h^{-1} Mpc)^3$  for  $k_{nl} = 64, 32, 16, 8,$  and  $4$  respectively. The purpose of these estimates is to give a rough idea of the range of parameters characterizing the models, and therefore more elaborate calculations are probably not needed.

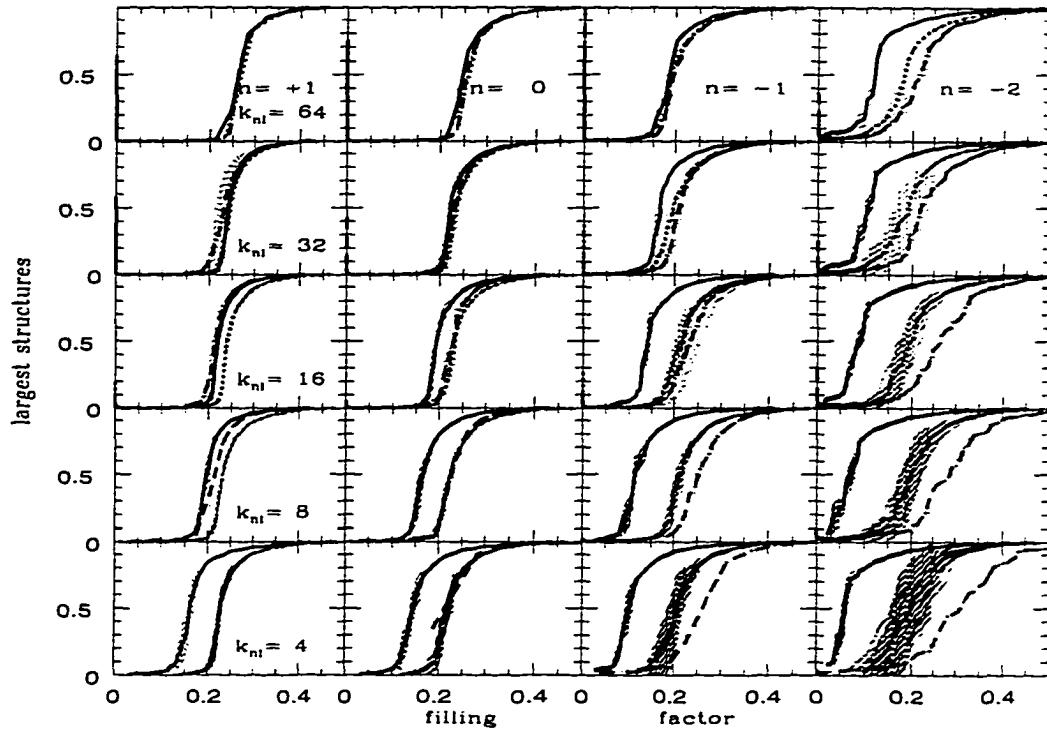


Figure 2.3: Percolation of nonlinear density fields derived from N-body simulations and the Gaussian randomizations. Cluster (over-dense) percolation is shown by solid lines, void (under-dense) percolation by dashed lines, and dots represent the percolation of the randomized fields. The volumes of the largest structures are plotted in units of the filling factor. Short dashes reflect dispersion at the  $1\sigma$  level.

## 2.5 Nonlinear Gravitational Distributions

### 2.5.1 Largest Cluster and Largest Void

Figure 2.3 shows the largest cluster and largest void statistic for families of four N-body models covering the complete range of initial power spectra taken at all five stages of the evolution. Each panel shows three curves: the largest cluster and void in the N-body simulation, and the largest structure in the Gaussian field having identical Fourier amplitudes with the N-body simulation. Again, in Gaussian fields there is no statistical difference between the largest cluster and largest void because of the symmetry of the distribution.

In Gaussian fields the percolation transition happens at a filling factor of about 16% corresponding to  $\nu = \pm 1$ . However, the finite size and resolution of the sample biases the transition. In order to avoid these effects, we obtain the "Gaussian" distribution by mixing the phases as described before. This allows for the generation of as many Gaussian realizations with identical amplitudes as needed to estimate the dispersion. The Gaussian, largest structure is shown as a dotted line in Figure 2.3 (hidden by the shade of the error bars) usually lying between the solid and dashed lines.

The  $n = +1$  model demonstrates the smallest difference between the properties of the largest cluster and largest void. Both percolate slightly, yet significantly, better than the corresponding Gaussian field. The  $n = 0$  model shows that the under-dense regions percolate similarly to the corresponding Gaussian field, but the over-dense regions percolate better than the corresponding Gaussian field. These models exhibit a network topology for both clusters and voids. In the  $n = +3$  model (not shown) the under-dense phase consistently percolates easier than the over-dense phase, and the over-dense phase percolates similarly to the Gaussian field except that it percolates easier in the last evolutionary stage ( $k_{nl} = 4$ ).

The major feature of the nonlinear distribution is that the largest cluster percolates easier than in the Gaussian case. The significance of this conclusion for the largest cluster is at the many- $\sigma$  level for most of the patterns. Qualitatively this remains true for other models we have studied earlier (CDM, C+HDM (Klypin & Shandarin 1993)), but quantitatively the transitions are different. The over-dense regions form a connected network spanning through the whole region when the filling factor is relatively small (smaller than in the Gaussian case), and this transition can be labeled as a shift toward a network structure. On the contrary, the under-dense regions may not always form a percolating void until the filling factor of the low density phase is significantly greater than that in the Gaussian

field. This type of transition can be labeled as a shift toward a bubble structure. The range of the filling factor corresponding to a sponge topology is typically (but not necessarily) increased compared to the Gaussian case. Thus, the above changes can also be labeled as a shift to a sponge topology. Results (not shown) from the percolation of the  $n = 3$  models demonstrate a smooth transition from a connected topology for both clusters and voids to the bubble topology of the  $n = -2$  models. In addition, the  $n = 3$  case supports our claim of the universality of a filamentary nature for the mass distributions. However, the major point is not how to label a structure but rather to show that in a general case the two shifts are independent of each other and carry independent information about the structure. Therefore, combining them into one parameter (like  $W_\nu = \nu_+ - \nu_-$  mentioned above) results in the loss of information. Similar to Gaussian fields at small filling factors, the largest structure must be negligible in sufficiently large samples to insure a fair sample.

### 2.5.2 Density Contrast of the Largest Cluster

The significance of the percolation transition in the nonlinear gravitational distributions sometimes is challenged on the grounds of the low density threshold of the percolation onset. However, these numbers may be misleading. We believe that the mean density is more relevant. Let us imagine that we have two identical density contours with  $\delta = 0.5$  in two dimensions. Within one of them there are a few peaks the highest of which is  $\delta \approx 1$  and within the other one there are few peaks of order  $\delta \approx 10$ . Now the question is which of these two contours will be picked up by the eye? We believe that obviously the latter is more noticeable. In the linear regime both the mean density and the density threshold are close, but in the nonlinear regime they are very different. Figure 2.4 shows the mean density of the largest cluster as a function of its volume given in units of the filling factor. For comparison, the density contrast is plotted as a thin line

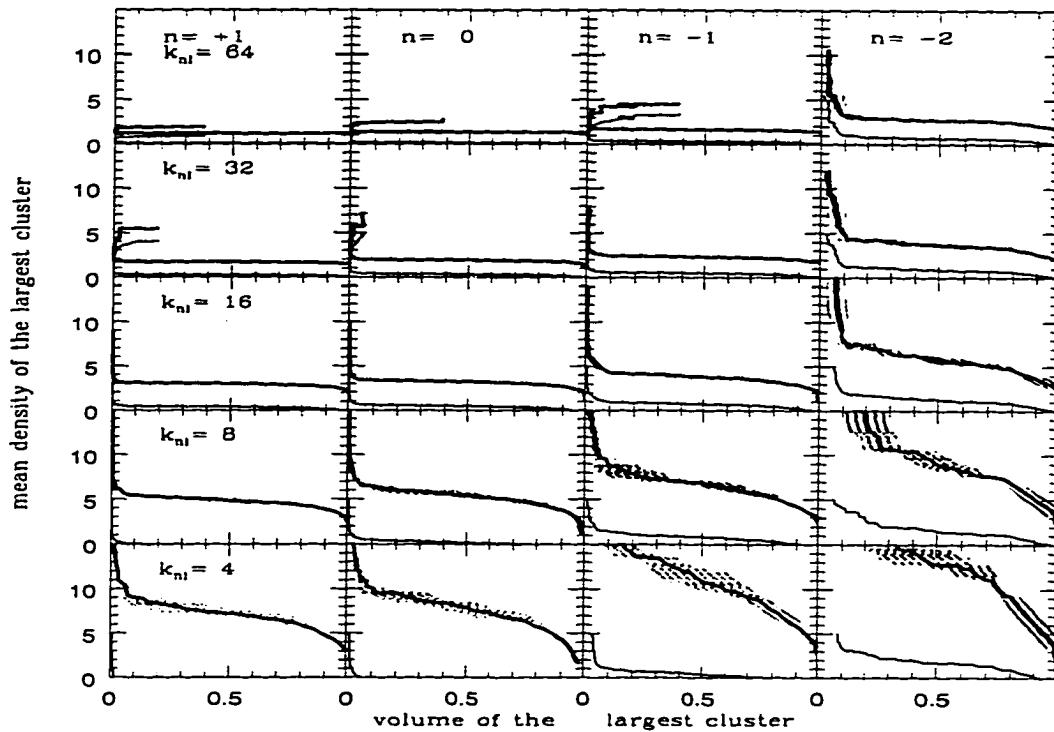


Figure 2.4: The mean densities of the largest cluster produced in N-body simulations compared to the normalized volumes. Where visible small dashes represent  $1\sigma$  dispersions. The thin lines are the value of the density contrast supplied for comparison.

and is markedly below the mean density in every case. An interesting feature of this figure is the stability of the mean density of the largest cluster through the percolation transition. The other important aspect of the figure is the value of the mean density at percolation. At percolation the mean density of the largest cluster is well above the average density of the distributions and understandably increases with evolution and decreasing spectral index. The double-valued nature of the function for the early stages of some models is an effect of a finite system explained above (section 2.2).

### 2.5.3 Number of Clusters and Voids

Another statistic characterizing the density distribution is the total number of clusters and voids as a function of the filling factor. As mentioned before this statistic is sensitive to the slope of the spectrum in the Gaussian fields. The number of clusters and voids is normalized to the Nyquist volume (8 mesh cells) as was done for Gaussian distributions. Frequency distribution studies establish that the smallest structures (both clusters and voids) dominate in the total number of structures for all spectra studied. Figure 2.5 shows this statistic for all models and all stages (except  $n = 3$ ). Three curves are plotted for each pattern. The total number of clusters and the total number of voids in the nonlinear distributions, and the total number of clusters (or voids) in corresponding Gaussian fields after randomization of phases. Typically the number of both clusters and voids in the parent N-body simulation is less than that in the corresponding nonlinear Gaussian field. Thus, the phase correlation due to nonlinear gravitational effects reduces the number of both over-dense and under-dense structures. The smaller the spectral index,  $n$ , the stronger the effect. The number of voids is consistently greater than the number of clusters in the nonlinear distributions if  $n \leq 0$ . On the other hand the number of clusters is greater in the  $n = +1$  model<sup>4</sup>. Another important trend is that at the later stages the differences between models become weaker compared to the beginning stages. With evolution, the models tend toward an estimated slope of  $n \approx -1$  in the nonlinear regime which is in very good agreement with the direct measurement of the spectra (Melott & Shandarin 1993).

Using the right middle panel of Figure 2.2 as a calibration, one can estimate the slope of the spectra for the nonlinear distributions by extrapolating between the maximum values from the pure power law, Gaussian fields. The maximum

---

<sup>4</sup>The last two stages ( $k_{nl} = 8$  and 4) in the  $n = +1$  model and the very last stage ( $k_{nl} = 4$ ) in the  $n = 0$  model suffer from discreteness. The voids are completely empty and percolate at such low density thresholds that the onset of percolation cannot be reliably calculated in the N-body simulations in question.

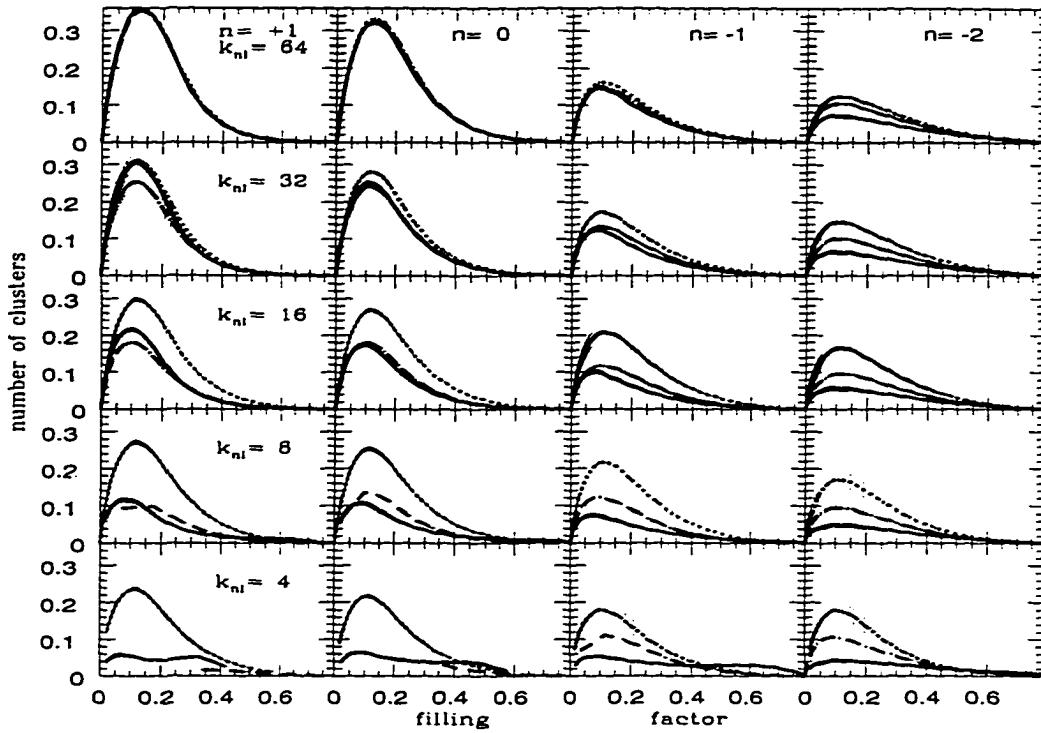


Figure 2.5: The number of clusters per Nyquist volume ( $\lambda_{Ny}^3$ ) for a given filling factor. Solid lines are statistics for clusters, dashed lines for voids, and dots for Gaussian randomizations. Short dashes represent  $1\sigma$  level dispersions.

values from both types of Gaussian fields are arranged in Table I with the estimated slopes in parentheses. Although the table is generally consistent with the trends described above, two anomalies are apparent in the results. First, the maximum value for the case  $n = -1$ ,  $k_{nl} = 64$  is considerably below the corresponding Gaussian value; and second, the models  $n = -1$  and  $n = -2$  do not monotonically approach a limiting slope of -1. Both of these peculiarities have a single explanation. Close inspection of the evolved power spectra for these models (see Melott & Shandarin 1993) shows that the same conditions apply to the slopes of the power spectrum if measured at  $k = 32$ . The Nyquist frequency of the Fourier transform used to produce the nonlinear Gaussian fields is  $k_{Ny} = 32$ . This suggests that our randomization procedure is sensitive to the slope of the

Table 2.1: Maxima of the Number of Clusters Statistic (Estimated Slopes)

n	Gaussian	N-Body Simulations				
		$k_{nl} = 64$	32	16	8	4
1	0.38	0.37 (0.4)	0.31 (-0.2)	0.30 (-0.4)	0.27 (-0.7)	0.24 (-1.1)
0	0.32	0.33 (0.1)	0.28 (-0.6)	0.27 (-0.7)	0.25 (-0.9)	0.22 (-1.2)
-1	0.24	0.16 (-1.8)	0.17 (-1.7)	0.21 (-1.3)	0.22 (-1.3)	0.18 (-1.6)
-2	0.14	0.12 (-2.2)	0.15 (-1.9)	0.17 (-1.8)	0.18 (-1.7)	0.15 (-1.9)

Maxima have errors of 10 percent.  
Estimated slopes are reported in significant figures.

power spectrum of the parent field at the maximum value of  $k$  used in the Fourier transform. In these cases, the density fields were  $64^3$  fields with maximum  $k$  values equal to the Nyquist wavenumber, but other cutoff values,  $k_c \leq k_{Nyq}$ , could be enforced. Further study is needed to make the dependence of the slope on the cutoff ( $k_c$ ) of the randomization procedure a practical technique to measure the effective slope of spectra. The models  $n = 1$  and  $n = 0$  do not exhibit the anomalies explained above because their power spectra correspond to the trends at the Nyquist frequency.

## 2.6 Discussion

The quantitative topology proved to be a useful measure for studying the large-scale structure in the universe. We use a percolation technique to study the mass distribution in real space obtained from N-body simulations of power law, initial spectra:  $P(k) \propto k^n$  with  $n = +3, +1, 0, -1$ , and  $-2$  in the  $\Omega = 1$  universe. Five stages corresponding to the scale of nonlinearity at  $k_{nl} = 64, 32, 16, 8$ , and  $4$  are analyzed. Each model was run with four different sets of random numbers to create ensemble families so that the dispersion of statistical parameters could be estimated.

We discuss in detail only two robust characteristics calculated in percolation

analysis: 1) the normalized volume of the largest cluster (or void), and 2) the total number of clusters (or voids) as a function of the filling factor. It is important to stress that the statistical properties of clusters and voids are identical in Gaussian fields, but they are significantly different in the nonlinear distributions drawn from the N-body simulations. Thus both statistics are sensitive to the non-Gaussianity of a field. These two functions can be very efficient discriminators of the models. All of the density distributions studied here have unique sets of the three functions (the filling factor, the largest structures, and the number of structures) on a multi-sigma level.

Combining the results that the most numerous clusters are the size of the cutoff scale, and that the nonlinear distributions percolate better than the Gaussian fields, we conclude that the small clusters are arranged in accordance with an underlying network structure. A similar conclusion was derived from the adhesion approximation (Kofman et al. 1992), and the truncated Zel'dovich approximation (Coles et al. 1993; Pauls & Melott 1995). The largest cluster as well as the largest void statistics indicate the percolation thresholds which are associated with a change in the topology of the distributions. The onset of percolation in the over-dense phase marks the transformation from a clumpy (meat-ball) to a network (sponge) topology, and the onset of percolation in the under-dense phase marks the transition from a bubble to a network topology. Comparing the simulations with the Gaussian fields we conclude:

- 1) For all models with  $n \leq +3$  over-dense regions percolate better than the corresponding Gaussian fields showing a shift toward a network structure. The later the stage the more significant the shift. Since the models are approximately scale free the later stages can also be interpreted as earlier stages of the same distribution seen with better resolution. Thus, we conclude that all power law models with  $n \leq +3$  show a network structure in the density distribution at the nonlinear stage if seen with sufficient resolution. Any differences are only

quantitative, but highly significant. The smaller the spectral index the stronger the shift. From this, we conclude that all realistic cosmological models (with non-power law, initial spectra like the CDM family) can be characterized as filamentary or network structures as far as the mass density in real space is concerned. At early stages of the CDM or C+HDM models when the effective slope at the scale of nonlinearity is small enough ( $n_{eff} < -1$ ), the density distribution is also of a bubble type.

2) The percolation transition happens at relatively low density contrasts. The larger the value of  $n$ , the smaller the density contrast at the percolation threshold (compare Figures 2.3 and 2.4). However, the density threshold is a misleading parameter. We propose the mean density of the largest cluster to characterize the prominence of the transition. The value of the mean density of the largest structure at the percolation transition indicates a definite nonlinear character for the largest cluster.

3) Under-dense regions percolate at lower filling factors than their Gaussian counterparts for  $n = +3$  and  $+1$ , and at considerably higher filling factors for  $n = -2$ . There is a smooth and consistent trend for under-dense regions to percolate with greater difficulty as the spectral index decreases or with evolution of a simulation. This trend represents a shift towards a bubble topology for models with  $n \leq -1$  indicated by the under-dense regions percolating at higher filling factor values than the corresponding Gaussian fields while the over-dense regions percolate at lower filling factors than the corresponding Gaussian fields. There is a noteworthy quantitative difference between the models  $n = -1$  and  $n = -2$ .

4) If the density threshold is selected so that the over-dense and under-dense phases each occupy 50% of the volume, then all models show that the largest cluster and void occupy almost all of the space. All smaller structures combined occupy at most a few percent of the volume for all the power law models studied.

Table 2.2: Volume Fractions in Clusters/Voids/Gauss (except the largest structure)

Index	$k_{nl} = 64$	32	16	8	4
$n = 0$	0.8/0.8/0.2	0.6/0.6/0.2	0.4/0.4/0.2	0.4/0.4/0.2	1.4/0.5/0.2
$n = -1$	0.8/0.8/0.5	0.5/0.6/0.5	0.5/0.7/0.5	0.4/0.8/0.4	1.0/1.0/0.5
$n = -2$	0.8/1.7/0.8	0.9/2.0/1.0	0.6/2.2/1.0	0.6/2.7/1.0	0.8/5.0/1.0

All fractions reported in per cent (at a filling factor of 0.5).

Values displayed are significant figures.

Thus, the distribution can be labeled as a sponge topology as predicted by Gott, Melott & Dickinson (1986). Table II gives the fractions of the total volume for all clusters and voids (except the largest structures) when the filling factor is 0.5, and the corresponding values for the Gaussian fields. Despite the smallness of the numbers they are highly significant. This explains why Weinberg, Gott & Melott (1987) found that the nature of the interlocking, equal-volume, over-dense and under-dense regions of a random field was a "sponge topology" where both the under-dense and over-dense regions are equivalent.

5) Our results show that all approximations for a nonlinear, gravitational instability based on nonlinear transformations of initial Gaussian density fields (e.g. the lognormal model) make incorrect predictions of the topology since they preserve the similarity in percolation properties of the over-dense and under-dense phases to that of the Gaussian fields.

# **Chapter 3**

## **Percolation Analysis of a Wiener Reconstruction of the IRAS 1.2 Jy Redshift Catalog**

### **3.1 Introduction**

Quantifying the distribution of galaxies in the visible universe has been one of the primary objectives of the study of the Large Scale Structure of the universe for several decades. With the compilation of early two dimensional galaxy catalogs (for example, Zwicky et al. 1961 and Shane & Wirtanen 1967), astronomers noted structures indicative of clustering and evolution. More recent redshift surveys (see. Davis et al. 1982; de Lapparent, Geller & Huchra 1986; Giovanelli & Hayes 1985; Tully & Fisher 1987; Lawrence et al. 1996) produced three-dimensional galaxy distributions and revealed structures such as voids (for example, Kirshner et al. 1981), and filaments and sheets of galaxies (Geller & Huchra 1989). Angular and spatial correlation functions (Peebles 1980 and references therein ) were employed as initial attempts to distinguish the galaxy surveys from random distributions and from comparisons with theoretical models.

In 1982, Zel'dovich proposed that the large scale structure of the universe could be characterized by its topology, and a multitude of statistical measures have been developed since then to quantify the topology of a distribution: the percolation threshold (Shandarin 1983; Shandarin & Zel'dovich 1983); the genus (Gott, Melott & Dickinson 1986); the contour crossing (Ryden 1988); random walk statistics (Baugh 1993); Minkowski functionals (Medke, Buchert & Wagner 1994); and minimal spanning tree characteristics (Bhavsar & Splinter 1996). Many obstacles have been overcome in the refinement of the measures above including boundary and selection effects, discreteness, local biasing, and error estimation. For example, recent topological analyses of the CfA redshift survey by percolation (de Lapparent, Geller & Huchra 1991) and genus (Vogley et al. 1994) report important and complementary findings. The topology and power spectrum of the IRAS survey has been examined in the context of the Queen Mary and Westfield College, Durham, Oxford and Toronto (QDOT) survey by Moore et al. (1992). The results of every report cited above are consistent with the findings of Gott et al. (1989): On scales significantly larger than the correlation length the topology is sponge-like. A sponge topology is characterized by equivalent, over- and underdense, multiply connected regions both of which percolate and are completely interlocked. These findings are consistent with the scenario that large scale structure developed from initial random Gaussian density fluctuations present at the epoch of recombination. In addition, for smoothing lengths comparable to or smaller than the correlation length slight shifts towards a 'meatball' (Gott et al. 1989; Moore et al. 1992) and 'bubble' (Vogley et al. 1994) topology have been reported. The scope of the upcoming Sloan Digital Sky Survey (Gunn & Weinberg 1995) promises even more significant results.

We use a percolation analysis that tracks two parameters in order to measure the connectivity of a Wiener reconstruction of the IRAS 1.2 Jy Redshift Survey and to estimate the spectral index of its associated power spectrum. The Wiener

filter (WF) takes a three dimensional, redshift map with the excluded zones filled by interpolation and converts it to a noise-free, full sky, real space map. Our percolation code computes the normalized volume of the largest structure as a function of the filling factor for a topological comparison. The number of clusters statistic has been shown to be sensitive to the index of the power spectrum for a simple power law relationship (Yess & Shandarin 1996). The method has been used for studying the properties of voids (Sahni, Sathyaprakash & Shandarin 1994) and the geometry of mass clumps (Sathyaprakash, Sahni & Shandarin 1996) in cosmological N-body simulations. The largest structure can be an over-dense region (cluster<sup>1</sup>) or an under-dense region (void). By comparing the growth of the largest structures (as functions of the filling factor) in a distribution with their growth in a randomized version of the distribution the distribution topology can be characterized. A Gaussian randomized version of a distribution is by definition a structureless field with the same power spectrum as its parent, so that any distribution that percolates at a lower filling factor than its randomized version is considered more connected than a random field, and hence, a connected network; whereas, any field that percolates at a higher filling factor than its randomized version is considered isolated or clumpy. In addition, reconstructions of N-body simulations (with and without Wiener filters) over the spectral range  $n = -2, -1$  and 0 are analyzed for comparison.

In §2 we describe the density fields derived from Wiener reconstructions of the IRAS 1.2 Jy Redshift Survey (Fisher et al. 1995a) and randomized versions which preserve the underlying power spectrum but are Gaussian fields. In §3 we detail the percolation method and parameters we have developed for analysis and comparison. Our results are presented and scrutinized in §4 with conclusions to

---

<sup>1</sup>The terms cluster and void in this context refer to high and low density regions respectively and not to the common astronomical meanings. Also, for the IRAS survey the terms under-dense and over-dense refer to galaxy densities; whereas, for N-body simulations the terms refer to mass densities.

follow in §5.

## 3.2 The Density Fields

The core of this study are Wiener reconstructions of the *real* space density field formed from the IRAS 1.2 Jy Survey (Fisher et al. 1995b). This survey is an extension of the 1.936 Jy flux limited survey of Strauss et al. (1992) compiled from the Infrared Astronomical Satellite Point Source Catalog (1988; PSC). The survey contains 5321 galaxies which cover 87.6 per cent of the sky. The 12.4 per cent of the sky that is missing from the survey is the Zone of Avoidance ( $\pm 5$  deg from the galactic plane) and a few confused regions or areas lacking coverage. For specifics of the galaxy selection criteria, the participating telescopes, data reduction techniques and results, the derived selection function, and the galaxy distribution see Fisher et al. (1995b).

### 3.2.1 Wiener Reconstructions

A Wiener reconstruction method is used to convert an interpolated redshift density field to a real space density field while suppressing noise. This approach is valid in the context of linear theory implying significant smoothing which is an aspect of the filter. The Wiener filter reconstruction method has been employed in many fields (Rybicki & Press 1992) to enhance a signal in the presence of noise. In cosmology, one use of the method is the minimum variance reconstruction of the real space density function from an incomplete and sparsely sampled galaxy distribution in redshift space. The process depends upon an expectation of the clustering properties of the real underlying density field being probed by the galaxy survey. In the cosmological case, the underlying density field is assumed to be Gaussian up to near the stage of non-linearity.

The WF algorithm as it applies to the IRAS 1.2 Jy galaxy survey is discussed

in detail in Fisher et al. 1995a; for completeness we give a brief summary here. The filtering is applied to a three dimensional decomposition of the redshift space galaxy density field in an orthogonal basis set of the spherical harmonics and spherical Bessel functions. The decomposition is truncated with  $l$  ranging from  $0 \leq l \leq 15$  (with  $-l \leq m \leq l$ ) and  $0 \leq k_n r \leq 100$  as a compromise between resolution and the number of expansion coefficients. The WF reconstruction technique depends on the assumed linear theory growth parameter  $\beta = \Omega^{0.6}/b$  where  $\Omega$  is the current cosmic density and  $b$  is the linear bias parameter. We investigate a set of ten reconstructions spanning the range  $\beta = 0.1$  to  $\beta = 1.0$ . In each case, the real space density field is reconstructed on a  $64^3$  grid with sides of length  $200 h^{-1}$  Mpc (20,000 km/s).

To compute the Wiener reconstructions the first step is to compute the redshift space harmonics in the spherical harmonics and spherical Bessel functions basis. This is analogous to computing the Fourier components in the analysis of the power spectrum. The redshift harmonics are distorted from the values that would be measured in a perfect real space galaxy distribution. First, the actual galaxy distribution is sparsely sampled and this results in a statistical uncertainty or shot noise in the estimated harmonics. Second, peculiar velocities introduce a systematic distortion due to the coherent infall and outflow around over-dense and under-dense regions. In linear theory, this redshift distortion can be computed if the value of  $\beta$  is known; the spherical basis function is convenient here since the distortion is in the form of a matrix which couples the radial modes of the expansion.

In the absence of shot noise, the real space harmonics could be recovered by a direct inversion of the coupling matrix. Shot noise makes this inversion highly unstable. The Wiener filter is a smoothing algorithm which is designed to make the inversion in the presence of noise optimal in the sense of minimum variance. It depends on the ‘prior’ which is the knowledge of the clustering of the

underlying field. Essentially, the Wiener filter is the ratio of the variance in the signal (determined from the assumed prior power spectrum) to the sum of the variance in both the signal and noise (determined by the amplitude of the shot noise).

The result is a density field in real space centered on the local group. Two important properties of the reconstruction method are an effective smoothing of the resultant field which increases with radius due to the limited resolution caused by truncating the harmonic expansion at  $l_{max} = 15$ , and the increased attenuation of the signal as a function of radius because the Wiener filter is dependent on the shot noise. The direct implication of truncating the harmonic expansion at  $l_{max}$  for fields sampled on a  $64^3$  grid is that at distances of  $R = 30$  mesh units ( $\approx 100 h^{-1}$  Mpc) the minimum resolution is approximately  $R \times (\pi/l_{max}) \approx 6$  mu. The effects of smoothing will be examined in detail in section §4. For detailed explanations of the Wiener filter reconstruction method for different response functions see Fisher et al. (1995a) and Zaroubi et al. (1995).

In addition to the reconstructions of the IRAS data, we produced reconstructions of cubic ( $L^3$ ) density fields derived from N-body simulations with power law initial spectrum ( $P(k) = 4k^n$ ) for  $n = -2, -1$ , and 0 evolved to the stage where scales of the size  $L/4$  were approaching nonlinearity (Melott & Shandarin 1993). Assuming that the rms fluctuation in the number of galaxies is approximately equal to the rms mass fluctuation, both are unity within spheres of radius  $8 h^{-1}$  Mpc. So, by identifying the stage where  $L/4$  becomes nonlinear with the present, a rough estimate of the size of a mesh cell is  $3.1 h^{-1}$  Mpc. These reconstructions were produced with and without Wiener filtering in order to systematically study the effects of harmonic expansion and Wiener filtering. In addition, the N-body simulations are reconstructed with the IRAS WF and not a WF based on the clustering and noise in the simulations. Using all the particles from the simulations to compute their harmonics and smoothing with a WF which corresponds to the

sampling density of the IRAS 1.2 Jy survey assures that the N-body reconstructions show the same resolution as the IRAS reconstructions. The volume of the N-body reconstructions was chosen to match the local mass density of the IRAS galaxy distribution at 500 km/s ( $5 h^{-1}$  Mpc), approximately  $0.045/h^{-3}$  Mpc<sup>3</sup> to allow for visual comparison.

### 3.2.2 Gaussian Randomizations

We produce Gaussian random fields by two methods in this study. Reconstructions are expanded in a spherical harmonic basis set and are also randomized in this basis; whereas, original N-body simulation density fields which are not reconstructed are Fourier transformed and then randomized in k-space. The crux of both processes is the randomization of the phases while retaining the original power spectrum of the parent field. For reconstructions (with and without Wiener filtering) this is accomplished by multiplying the density,  $\rho_{lmn}$ , for  $m > 0$  by  $e^{i\phi}$ , where  $\phi$  is a random variable in the range  $0 \leq \phi \leq 2\pi$ . The  $m = 0$  term is multiplied by  $\sqrt{2} \cos \phi$ , and the  $m < 0$  terms are determined using a reality condition for  $\rho$ . The randomized versions of fields derived from N-body simulations are created by multiplying the components of all k-space vectors by  $\cos \phi$ , and a reality condition again assigns values to coefficients in the lower half of k-space.

## 3.3 Percolation Method

The percolation methodology we employ analyzes galaxy (IRAS) or mass (N-body) distributions as well as void distributions. The intent is to characterize the topology of both distributions, and to estimate the slope of the power spectrum of the density field. The discriminator between mass sites and voids is the density threshold, and it is smoothly varied to establish contours separating clusters and voids. Void and galaxy percolation are analogous to mass percolation so for

simplicity percolation will be discussed in the context of mass percolation except where distinction is needed. As the density threshold is varied three parameters are tracked: the filling factor, the volume of the largest structure (for both over-dense and under-dense structures), and the number of isolated structures.

The filling factor is the fractional volume of all mass sites identified in the distribution for a given density threshold. It is equivalent to the cumulative distribution function for the clusters and the volume fraction of Gott et al. (1989) for Gaussian distributions. For clusters, the filling factor grows from a minimum value of zero to a maximum of one as the density threshold is systematically lowered. The filling factor serves as the independent variable for our functions to allow for a fair comparison between different density fields.

The second parameter, the volume of the largest cluster, is a stable indicator of the percolation transition and is used to assess the topology of the field. The volume is reported in units of the filling factor (the ratio of the largest cluster to the total volume of all clusters) as a function of the filling factor, and a rapid increase in the volume indicates the filling factor associated with the percolation transition. This transition represents a change from a clumpy to a connected topology for the field. For clusters it is a change from a meatball to a sponge topology, and for voids it is a change from a bubble to a sponge topology. Gaussian fields are used as standards of comparison to characterize the topology of density fields. A field which percolates at a smaller filling factor than its Gaussian counterpart displays a shift towards a connected topology, while a field that percolates at higher filling factors displays a shift towards a clumpy topology.

The number of clusters statistic is sensitive to the slope of the power spectrum of a field (Yess & Shandarin 1996). This implies that for a field described by a power spectrum of the form  $P(k) = Ak^n$  that the number of clusters statistic is sensitive to the spectral index,  $n$ . In fact, the maximum of the statistic is a function of  $n$ , and can be used to estimate the effective spectral index of a mass

distribution.

### 3.4 Results

Topological analysis of modern redshift surveys has focused on the two aspects of the galaxy distributions mentioned above: a quantitative assessment of the connectedness of the structure, and the slope of the power spectrum. In addition to the difficulties of assessing boundary effects and error estimations, a major obstacle for all current methods employed to describe the spatial distribution of galaxies is the lack of resolution resulting from the sparse sampling achieved in existing surveys.

The resolution of any representation of a galaxy survey is ultimately a function of the mean galaxy density of the survey and the chosen smoothing method. In this respect the IRAS 1.2 Jy survey presents good prospects with the average galaxy number density higher than the QDOT survey value in the region  $R \approx 100 h^{-1} \text{ Mpc}$ . However, different groups have utilized different smoothing routines to produce density fields from the galaxy distributions of the surveys. For example, the smoothing methods of Moore et al. (1992) in their analysis of the QDOT survey are typical, but differ significantly from those utilized in the spherical harmonic reconstruction of our data. Moore and collaborators used a constant Gaussian filtering width determined by the inter-galactic spacing at the edge of the QDOT survey,  $\lambda = [S(r_{max})]^{-1/3}$ , where  $S(r)$  is the radial selection function. In a magnitude limited sample this choice ensures that the density field is not under-sampled while providing an unprecedented number of resolution elements for the QDOT survey. In contrast, the spherical harmonic reconstruction of the IRAS 1.2 Jy survey implies a variable smoothing with radius due to the finite cutoff of  $l$  in the spherical harmonic expansion. In addition, the Wiener filter suppressed the amplitude of the field as a function radius to mitigate the effects

of increasing shot noise (as determined by the selection function). The effect of the variable smoothing in the density field is evident in the results reported below.

Like all statistical measures our parameters are sensitive to the resolution and number density of the data, but they are relatively robust with respect to boundary effects and scale. Since the volume of the largest structure is normalized to the filling factor, and the number of clusters statistic can be normalized to the volume of the survey, the geometry and size of the survey does not determine their analytic behavior. We exploit the stability of our parameters by analyzing spherical subregions of the survey in order to examine the effects of variable smoothing and any local bias against a fair sample. A measure of the stability of our parameters are the errors presented for the results from N-body simulations with multiple realizations. In all instances the error bars represent  $1\sigma$  deviations over four realizations. We do not estimate errors in the results for the IRAS reconstructions, but rather rely on trends in the versions varying in  $\beta$ , over the range  $0.1 \leq \beta \leq 1.0$ , to determine conclusions.

### 3.4.1 Largest Structures

Largest structure results for all versions of the Wiener Reconstructions are shown in Figure 3.1 for both clusters and voids. The top panels show the growth of the largest structures for a field sampled on a cubic grid,  $64^3$  mu ( $200 h^{-1}$  Mpc) to a side<sup>2</sup>. If we consider the percolation threshold to be the first significant jump in the value of the largest structure statistic (de Lapparent, Geller & Huchra 1986) then percolation happens for all versions in the range  $0.024 \leq ff \leq 0.05$  for clusters and between  $0.01 \leq ff \leq 0.022$  for voids. For cluster percolation, fields with larger values of  $\beta$  percolate at smaller filling factor values, while for voids the trend is generally reversed. This means the larger the average cosmic density or

---

<sup>2</sup>The result of a spherical harmonic reconstruction is a spherical field, and the ( $64^3$  m.u.) density field analyzed is the largest cubic subregion of the original output.

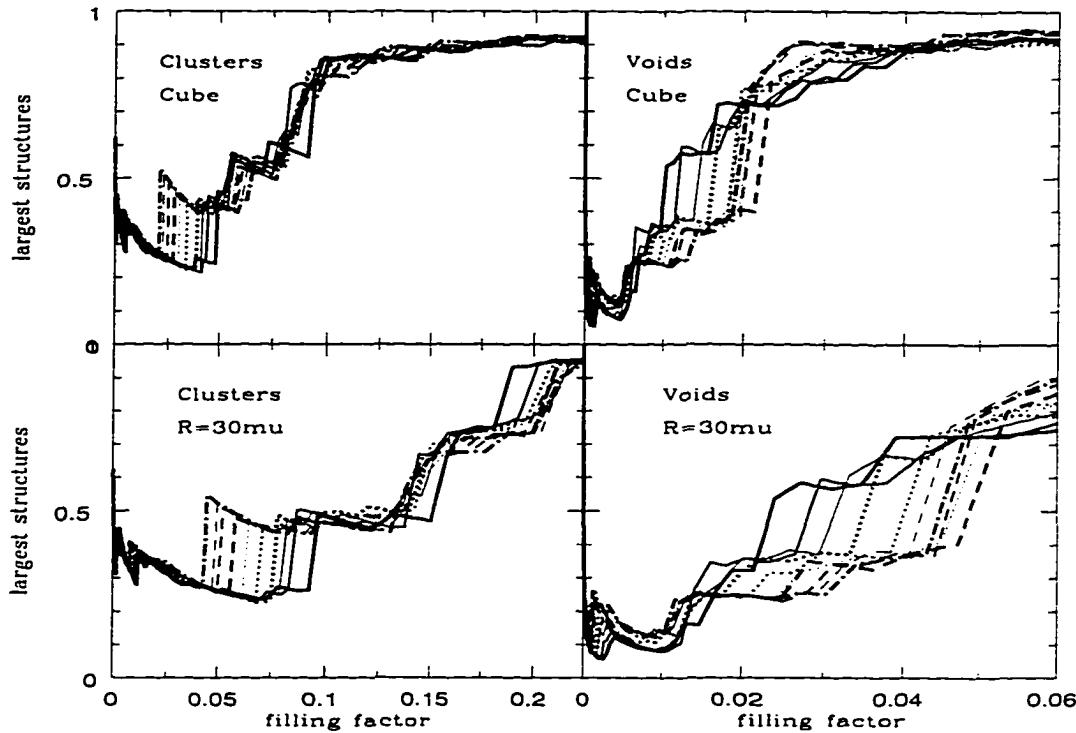


Figure 3.1: Largest Structure Statistic for IRAS reconstructions over the range  $0.1 \leq \beta \leq 1.0$ , where  $\beta = \Omega^{0.6}/b$ . Lines are in triplets, except  $\beta = 1.0$ , with lighter lines representing larger  $\beta$ : solid ( $\beta = 0.1, 0.2, 0.3$ ), dots ( $0.4, 0.5, 0.6$ ), dash ( $0.7, 0.8, 0.9$ ) and dot-dash ( $1.0$ ). The top panels correspond to the measurements in a cube of length  $200 h^{-1} \text{ Mpc} = 64$  mesh units on a side, and the bottom panels to measurements in a sphere  $30 \mu\text{m}$  in radius.

the smaller the bias factor the more connected the clusters tend to be. The fact that the percolation threshold is the distinguishing difference between the curves demonstrates the sensitivity of this parameter as suggested by Shandarin (1983). Alternately, the high sensitivity may also cause problems in noisy samples as reported by Dekel & West (1985). However, rigorously determining a percolation threshold value is not important to the analysis in this study and is used here only for illustration. In this study the shape of the largest structure function over its entire range will be used as a comparison to characterize the topology of a field.

Another important feature of the largest structure statistic is the high initial

values at low filling factors for all versions of the Wiener Reconstructions. This indicates that the largest structure is always dominant which indicates a problem with the size of the sample. In a statistically fair sample, there would be a multitude of small clusters at the high density cutoffs beginning the percolation process, and the largest cluster would emerge from the field as the percolation process caused clusters to join together. The fact that the statistic has a non-vanishing initial value is indicative of the relatively small sample size of this survey for the purpose of this statistic and the level of smoothing introduced by the reconstruction process. For comparison, see the percolation results of N-body simulations in Figure 3.3 and reference Yess & Shandarin (1996).

A lack of resolution at large distances explains the relationship between the results of the upper and lower panels of Figure 3.1. The growth of the largest structure function is virtually identical for the two cases except for a near doubling in the filling factor. This means that all information about structure is contained in the reduced spherical region bounded by  $R \approx 100 h^{-1} \text{ Mpc}$  (30 mu), and that the excess volume in the cube does not affect the structures but only contributes to a reduction in the filling factor. The implication is that the value of the field outside  $R \approx 100 h^{-1} \text{ Mpc}$  is featureless and roughly equal to the mean density. This is because attenuation of the signal is a function of radius due to the effective smoothing of the spherical harmonic reconstruction and the loss of detail after Wiener filtering. The effects of these two operations will be examined separately below.

Restricting the analysis to an even smaller volume ( $R \approx 30 h^{-1} \text{ Mpc}$  (10 mu), not shown), reveals similar percolation curves to those of larger volumes for the growth of the largest clusters except that volume effects for clusters are exaggerated to the point where the largest cluster is associated with the highest density peak and its volume is never less than half the volume of all clusters combined. For voids the curves are also similar at the outset, but rise much slower so  $ff \approx 0.3$

when the volume of the largest void approaches unity.

In Figure 3.2, we display the results of a systematic study of the effects of spherical harmonic reconstruction and Wiener filtering separately on N-body simulations. The results of analysis of four realizations of N-body simulations characterized by an initial power law power spectrum of the form  $P(k) \propto k^{-1}$ , evolved to the stage where  $\lambda = \lambda_f/8$  (where  $\lambda_f$  is the fundamental wavelength) is approaching non-linearity<sup>3</sup>. The upper panels show the results of percolation analysis for the simulations and demonstrate that the topology of the structure is similar throughout the volume and not a function of radius. The difference between percolation in a cube and a sphere is also insignificant. It is also important to note that the rapid growth of the largest structure for random Gaussian fields (light lines) starts at a filling factor of  $ff \approx 0.16$  for both clusters and voids, which is the expected value. The interpretation of this data is that the topology of these simulations is characterized by a very connected cluster network and slightly more isolated voids compared to Gaussian fields.

The middle panels show the effect of spherical harmonic reconstruction alone and in conjunction with Wiener filtering. The reconstruction process eliminates the distinction between the topologies of clusters, voids and the random fields in a spherical volume of radius,  $R \approx 100 h^{-1} \text{ Mpc}$ , and introduces some distortion in the curves at low filling factors. In addition, the Wiener filter removes most of the small scale structure demonstrated by the almost immediate percolation of the fields and the reduction of error bars. It is easy to understand this effect of the Wiener filter because it reduces the range of the density values in the reconstructed density field by one-fourth. In order to regain a measure similar to that of the original density fields, the effects of the reconstruction and filtering have to be minimized by restricting the analysis to smaller radii. The smoothing effect of

---

<sup>3</sup>For a detailed discussion of the N-body simulations used in this study see Melott & Shandarin (1993).

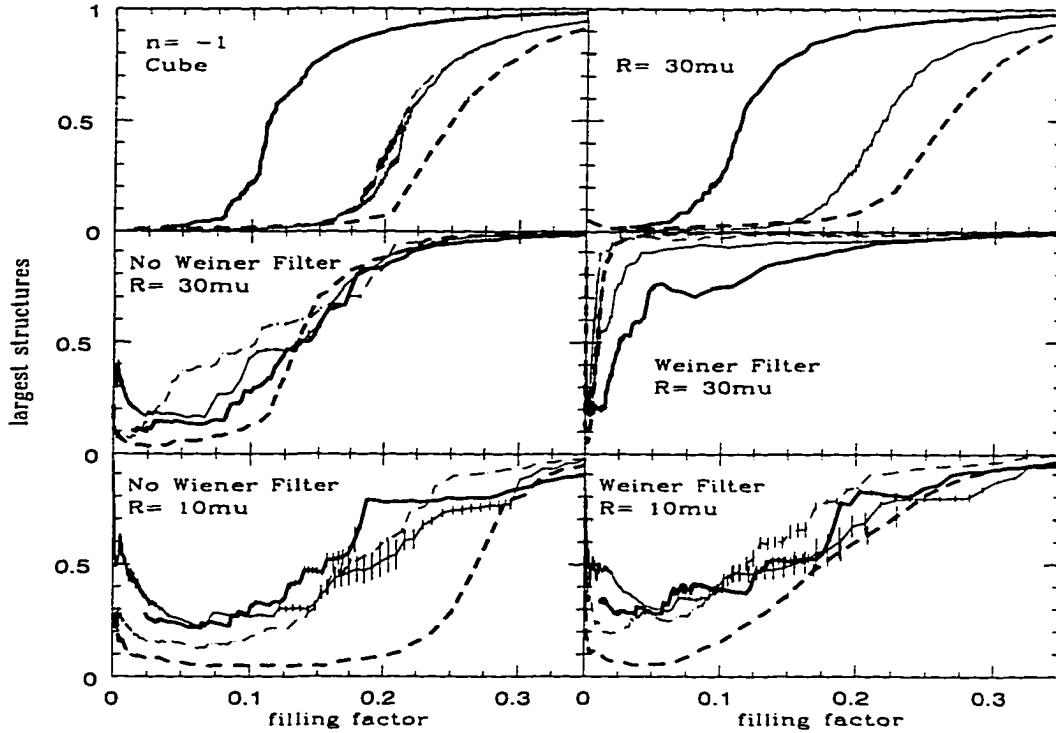


Figure 3.2: Largest Structure Statistic for N-body simulations described initially by a spectral index  $n = -1$  under a variety of conditions. The top panels report results from the analysis of density fields derived from the initial N-body simulations; whereas, the middle and lower panels display the results from analysis of reconstructions of the density fields. Solid lines indicate cluster (over-dense) percolation while dashed lines indicate void (under-dense) percolation. Light lines represent Gaussian randomizations. Error bars are  $1\sigma$  deviations over four realizations in all cases, although commonly they are smaller than the line thickness.

the reconstruction is less and the signal to noise ratio is better for smaller radii, so that the topology of the original field can be recognized in a volume limited sample at  $R \approx 30 h^{-1} \text{ Mpc}$  if the reconstruction alone is applied to the simulation (bottom left panel). The Wiener filtering still distorts the original topology even in this restricted volume as shown in the bottom right panel. An important feature for the interpretation of percolation results is that the largest structures in the random Gaussian realizations behave, making allowance for survey volume effects and smoothing, generally as expected in all cases except the middle right panel.

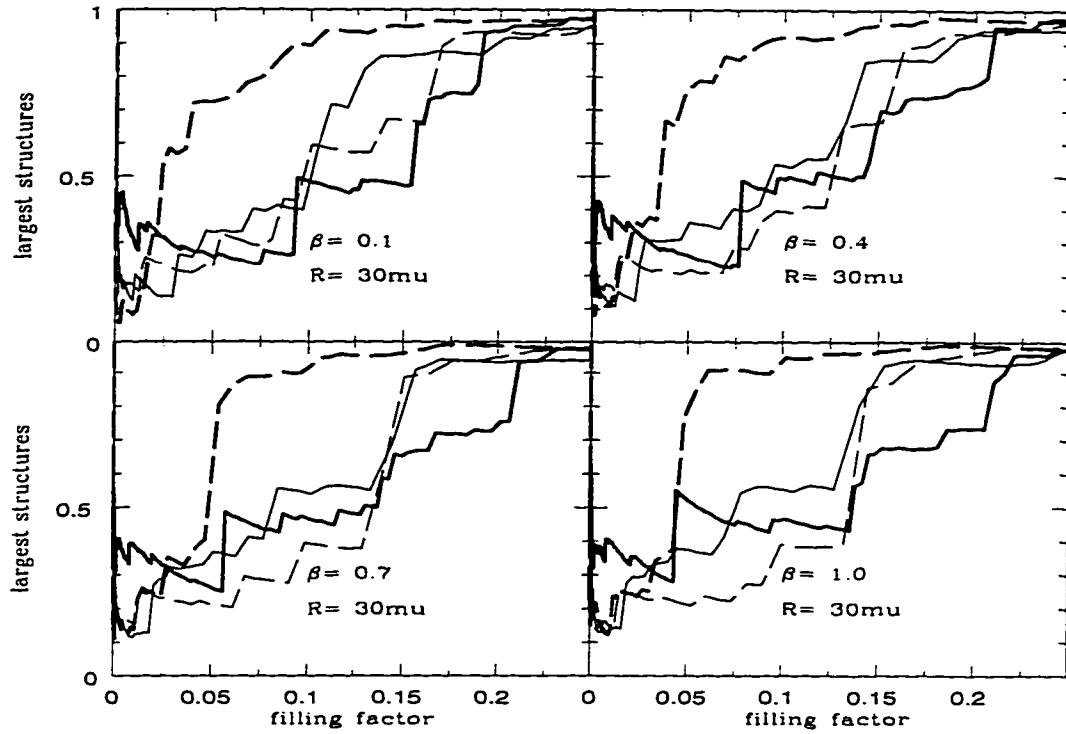


Figure 3.3: Largest Structure Statistic for selected IRAS reconstructions including Gaussian randomizations. Solid lines indicate cluster percolation while dashed lines indicate void percolation. Light lines represent Gaussian randomizations.

Finally, Figure 3.3 shows the largest structure statistics for Wiener reconstructions of the IRAS 1.2 Jy Survey for various  $\beta$  values. Each version displays a similar result with voids percolating at lower filling factors than the Gaussian counterparts, and clusters percolating similarly to the Gaussian fields. These results imply a well connected void distribution with a generally sponge like or slightly meatball cluster distribution. Volume and smoothing effects are again evident in each field demonstrated by the high values of the largest structure statistic at low filling factor values. This problem is more apparent in the cluster analysis of the original IRAS reconstructions than for voids or clusters in randomized IRAS reconstructions. Although the lack of resolution prevents a strong characterization of the topologies represented in the data, the results are consistent with

the slight meatball topology shift reported by Moore et al. (1992) and Gott et al. (1989) for similar local volumes. Results for volumes with  $R \approx 30 h^{-1} \text{ Mpc}$  (not shown) are inconclusive.

### 3.4.2 Number of Clusters

It has been established that the maximum of the number of clusters statistic reflects the slope of the power spectrum for density fields described by a simple power law (Yess & Shandarin 1996). This is true for Gaussian fields and randomized density fields derived from N-body simulations over a wide range of evolutionary stages. This statistic can easily be normalized by the volume of the field so that different samples can be directly compared: however, the values of the maximum are small enough in this study that the raw data is presented for clarity. The results of percolation analysis of clusters and voids are presented in Figure 3.4 for both reconstructed IRAS survey ( $\beta = 0.1, 0.5, 1.0$ ) and N-body simulation ( $n = 0, -1, -2$ ) fields. The most notable feature of the data is that the maxima for N-body simulation fields are two orders of magnitude below those of the original fields before they were reconstructed. This reduction in the number of clusters (voids) is a direct result of the smoothing and attenuation of the Wiener Reconstruction procedure. The effect is to reduce the signal below the level where distinctions can be made between fields characterized by different spectra. Another indication that the resolution of the reconstructions is insufficient for percolation analysis are the many local maxima in the statistic.

## 3.5 Conclusions

The Wiener reconstruction technique has proven successful in reconstructing the angular density fields of galaxies (Lahav et al. 1994); the temperature fluctuations of the Cosmic Microwave Background (Bunn et al. 1994); real space density,

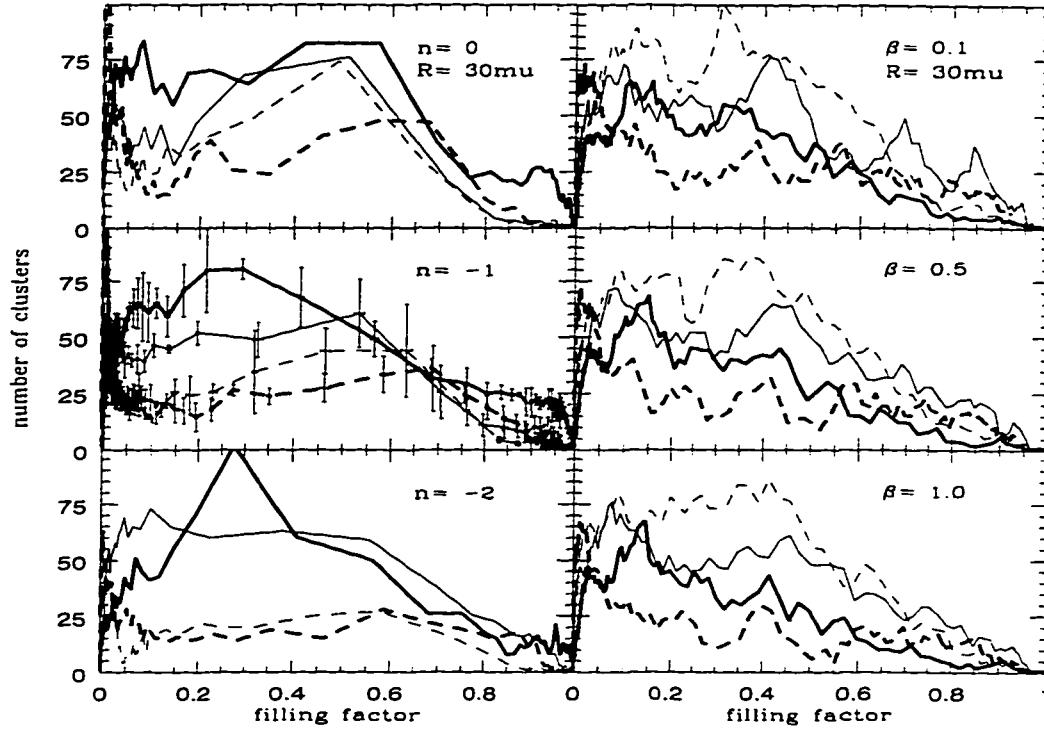


Figure 3.4: Number of Clusters Statistic for Wiener Reconstructions of N-body simulations ( $n = 0, -1, -2$ ) and IRAS data ( $\beta = 0.1, 0.5, 1.0$ ). Solid lines represent clusters and dashed lines represent voids, where light lines indicate Gaussian randomizations. Error bars for the  $n = -1$  case represent  $1\sigma$  deviations over four realizations.

velocity and gravitational potential fields (Fisher et al. 1995a); and predicted full sky density fields (Zaroubi et al. 1995). In this study we apply percolation analysis to full sky Wiener Reconstructions of the IRAS 1.2 Jy Redshift Survey. We find that our results are consistent with the conclusions of other studies that report a small shift towards a meatball topology for the survey region, however we would like to stress that our analysis was in real space.

The Wiener reconstruction technique smoothes the density field based on the inter-galaxy separation as a function of radius and attenuates the high frequency components or small scale components resulting from shot noise. This results in a loss of resolution with distance which presents a challenge for percolation

analysis. The largest structure statistic is robust enough to give an indication of the topology of the field under these conditions; however, the number of clusters statistic suffers too much from the loss of resolution to give a measure of the slope of the associated power spectrum. Alternately, the number of clusters statistic has the potential to be developed into an indicator of whether or not the structures of a given field represent a fair sample for statistical purposes.

The prospects of analyzing a full sky density reconstruction in order to assess the topology of the large scale structure and associate that structure with initial fluctuations in the matter density field at the time of recombination are attractive. This study offers an optimistic picture that as more galaxies are added to surveys the statistical measures presented will produce accurate and convincing results. Until galaxy survey counts are increased enough to overcome the problems identified in this study, percolation can still be applied to point-wise galaxy distributions, density fields of highly sampled portions of the sky or ideally volume limited subsamples.

# **Chapter 4**

## **Detection of Structure in the Las Campanas Redshift Survey**

### **4.1 Introduction**

For decades cosmologists have been developing methods for characterizing and quantifying the structure in the local galaxy distribution as supplied to them by astronomers. Numerous statistics have been employed and refined in this endeavor with three of the most successful being percolation analysis (Shandarin & Zel'dovich 1983; Einasto et al. 1984; Sahni, Sathyaprakash & Shandarin 1994), the genus statistic (Gott, Melott & Dickinson 1986) and minimal spanning tree (Bhavsar & Splinter 1996). By differing techniques, these statistics have produced compatible results describing the structure of the local universe in the IRAS 1.2 Jy Survey (Yess, Shandarin & Fisher 1997; Protogeros & Weinberg 1997). The minimal spanning tree statistic (MST) has been shown to duplicate the results of percolation analysis, but does not supersede it in descriptive capability. Both percolation analysis and MST differ from the genus statistic in that they are sensitive to geometry as well as topology. In topological studies, a major limiting factor has been the shot noise in the analysis of pointwise distributions, or resolution in the

analysis of density fields derived from galaxy positions. The completeness of the Las Campanas Redshift Survey (LCRS) in terms of galaxy numbers overcomes these discreteness effects. Also, the area of sky coverage in the survey promises that a fair sample of the universe is being probed. The extent of the upcoming Sloan Digital Sky Survey (Gunn & Weinberg 1995) promises equally unequivocal results over even larger regions.

The particulars of the Las Campanas Redshift Survey and our utilization of the survey are detailed in § 2 of this paper. In addition, the standard Poisson distributions and their application are explained in § 2. The parameters used to characterize the topology of the distributions are described in § 3 along with the percolation method for pointwise distributions. In § 4 the percolation results are presented and explained. Conclusions are also drawn in § 4 with suggestions for further investigations.

## 4.2 The LCRS and Poisson Standards

The Las Campanas Redshift Survey has two important attributes which overcome two of the most important problems with topological analysis, sparse samples and fair survey volumes. If a survey is too sparse in its sampling of the true galaxy population by design or because of magnitude limitations, the observed distribution is dominated by shot noise and the resulting topological determination is essentially that of a random distribution. There are approximately 25,000 galaxies with redshift positions in the LCRS. They are distributed over six slices, three northern and three southern. The slices, depicted in Figure 4.1, are strips of the sky  $1.5^\circ$  thick and  $80^\circ$  wide which are separated by  $3^\circ$ . The northern slices are centered at declinations of  $-3^\circ$  (N1),  $-6^\circ$  (N2), and  $-12^\circ$  (N3) and the southern slices at  $-39^\circ$  (S1),  $-42^\circ$  (S2), and  $-45^\circ$  (S3). All slices are probed to a depth of  $60000 \text{ km s}^{-1}$  ( $600 h^{-1} \text{ Mpc}$  for  $H_0 = h \text{ km s}^{-1} \text{ Mpc}^{-1}$ ) for galaxies of  $m = 17.75$ .

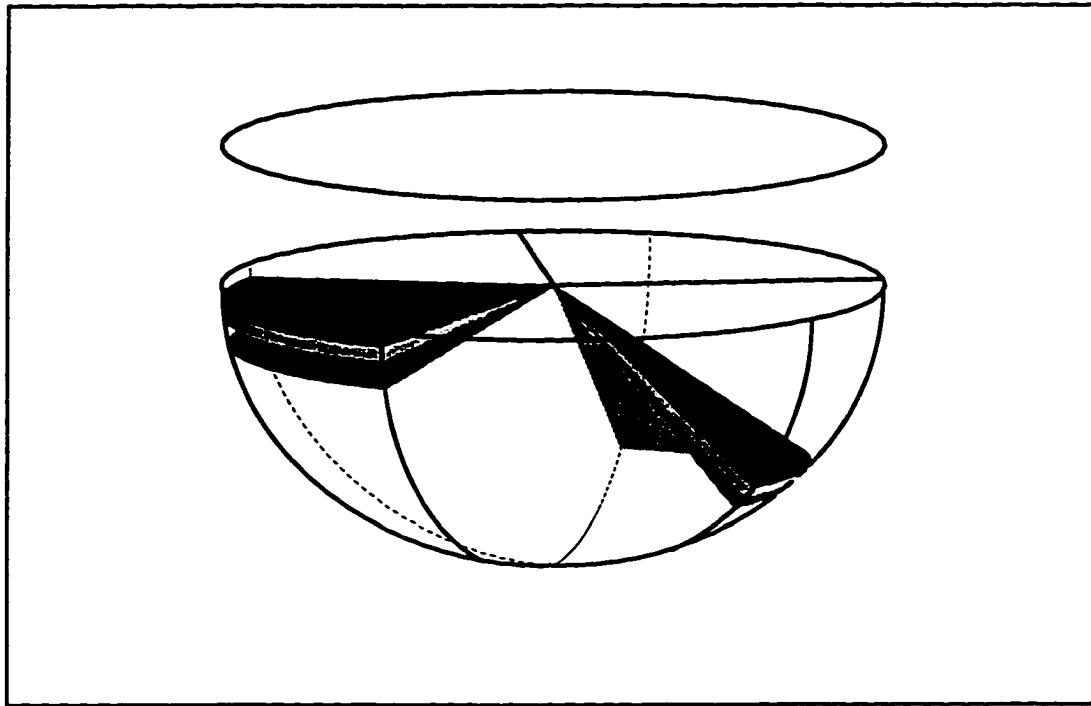


Figure 4.1: A representation of the Las Campanas Redshift Survey Slices and the two dimensional sectors onto which the galaxy positions are projected.

the limiting magnitude. For the details of the LCRS and clustering properties see Shectman et al. (1996) and references therein. The survey volume of the LCRS allows for a fair assessment of the topology given that cosmic structures are on the scale of  $100 \text{ h}^{-1} \text{ Mpc}$  and the number of galaxies contained in the survey gives a signal that, given the sensitivity of percolation analysis, overcomes random noise for a large portion of the survey volume.

In order to characterize the topologies of the Las Campanas' slices, standards typifying random distributions need to be constructed. Since the galaxy positions of the slices are projected to a central plane, the standards have to account for the local galaxy density of the LCRS plus the projection effects. In Figure 4.1, the consequences of projecting the galaxy positions to a plane are easily seen. The sectors at the top of the figure represent the two-dimensional, analysis areas.

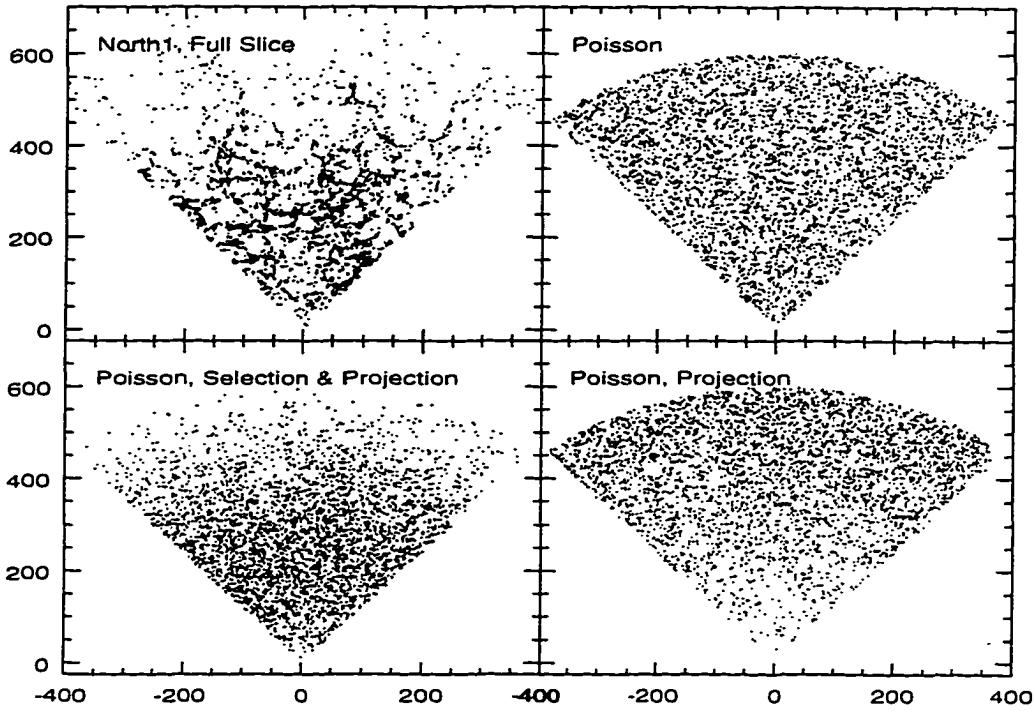


Figure 4.2: Map of the Las Campanas Redshift Survey Slice, North 1. There are 4495 galaxies plotted (initial Filling Factor of 0.016). To the side is the corresponding Poisson Distribution and below the distributions corrected for projection and selection and projection effects sequentially. The axis units are Mpc.

Poisson distributions created to correspond to an appropriate selection function and corrected for projection effects serve as standards for pointwise distributions. The selection function chosen to approximate the distributions of the six slices was taken from Lin et al. (1996) and based on the subsample of galaxies from both north and south slices termed NS112. Figure 4.2 shows the two-dimensional map of a northern slice, N1, of the LCRS in the upper left panel. Its corresponding Poisson distributions, with corrections sequentially applied, are also shown to illustrate the effect of each correction. Note, it is easy for the eye to discern the survey distribution from the random distributions.

### 4.3 Percolation

The first step in the percolation of pointwise distributions is to superimpose a grid on the sector geometry and then locate the galaxy positions on the lattice. In projecting the galaxy positions to a flattened, central plane, we have essentially ignored the effects of curvature associated with the geometry of the initial survey which may possibly have a greater consequence for the southern slices. Our percolation analysis is performed on a two dimensional lattice of cells  $1 \text{ Mpc}^2$  in area and the same size for all slices and standards. The positions of galaxies are equated with filled lattice cells. Filled cells that share a common side are considered neighbors. Through the stipulation that ‘any neighbor of my neighbor is my neighbor’ clusters<sup>1</sup> composed of adjacent cells are defined and grow.

Our percolation method for pointwise distributions allows for two means by which clusters can grow. Circles of specified radius are constructed around the initially filled cells (galaxy positions) in the distribution. The radii of these circles are incrementally increased to encompass adjacent sites. Cells enveloped by expanding circles are labeled filled and are considered neighbors of the initial cell at the center of the circle. If two or more circles come to overlap while expanding, the members of the overlapping clusters merge into a single combined cluster. As the radii increase clusters will grow in size and generally diminish in number due to mergers. This process will continue until the largest cluster is the only cluster in the distribution. In an infinite space the largest cluster emerges as the infinite cluster. For details of the pointwise percolation method see Klypin & Shandarin (1993).

In this study, we track two percolation parameters as functions of the increasing circle radius. The filling factor is defined as the fraction of filled cells in the total area. The second parameter is the relative size of the largest cluster to the total

---

<sup>1</sup>The term cluster as used in percolation analysis does not imply cluster of galaxies in the astronomical sense.

area of filled cells. This descriptor is reported in units of the filling factor so that its initial value should be small when the largest cluster is one of many clusters, and its maximum value is 1.0 when the largest cluster spans the space and incorporates all the filled cells of the distribution. The relative area of the largest cluster is reported as a function of the filling factor in comparisons between galaxy and Poisson distributions (see Figure 4.3).

A rapid rise in the largest cluster statistic is indicative of the percolation condition (Klypin & Shandarin 1993). However, it is not important for this study to determine the exact filling factor associated with the onset of percolation. The exact filling factor is a noisier statistic than the largest cluster statistic and less discriminating. In our method, the largest cluster statistic, over its full range, is used to characterize the nature of the distribution (Yess & Shandarin 1996). In general, the faster the largest cluster statistic grows, the more the curve shifts to the left and the more connected the distribution. A distribution for which the largest cluster statistic grows more rapidly than a comparable Poisson model is described as an example of a network topology and a distribution for which the largest cluster statistic grows more slowly is considered to be clumpy or have a ‘meatball’ topology.

For a direct comparison of percolation results, pointwise distribution standards need to have equivalent initial filling factors as the distributions they are characterizing. If the initial filling factor of a standard distribution is too high ( $ff_0 \geq 0.1$ ), the resolution of the percolation parameters will not be sufficient to detect the onset of percolation. The number of galaxies in the original slices and the initial filling factors of all the distributions analyzed are given in Table 1. The Poisson standards in this study are random distributions adjusted, as a function of radius, for selection and projection effects with corresponding initial filling factors well below the resolution limit. In addition, it is easy to see that because of our percolation method the resolution of the largest cluster statistic can depend

Table 4.1: LCRS Slice Data

slice	Full Slice		Magnitude	Volume
	dec	N	$ff_o$	$ff_o$
N1	$-3^\circ$	4495	0.016	0.029
N2	$-6^\circ$	2589	0.009	0.018
N3	$-12^\circ$	5217	0.018	0.030
S1	$-39^\circ$	4339	0.015	0.025
S2	$-42^\circ$	4728	0.017	0.027
S3	$-45^\circ$	4140	0.015	0.024

strongly on the value of the initial filling factor. For instance, for distributions that initially have well isolated galaxies the majority of the clusters will contain only one filled cell. In this case, the first iteration of the expanding circles will add four nearest neighbors to virtually every cluster causing the filling factor to increase by a factor of nearly five. The initial filling factors of the Las Campanas slices are well below the level where this effect would negate the results.

#### 4.4 Results and Conclusions

Shown in Figure 4.3 are the results of the Las Campanas percolation analysis. The left column shows the results for the three northern slices of the survey while the right column shows the results of the southern slices. The solid lines are the results for the survey slices with the lightest being N1 (S1) and the heaviest N3 (S3). In all graphs, the dotted line to the right is the result from a pure Poisson distribution with the appropriate initial filling factor and geometry. This result is shown for reference. The dashed line is the result from an appropriate Poisson distribution corrected for selection and projection effects. In all cases the error bars represent one sigma deviations over four realizations.

In the top panels the corrected Poisson results are indistinguishable from the

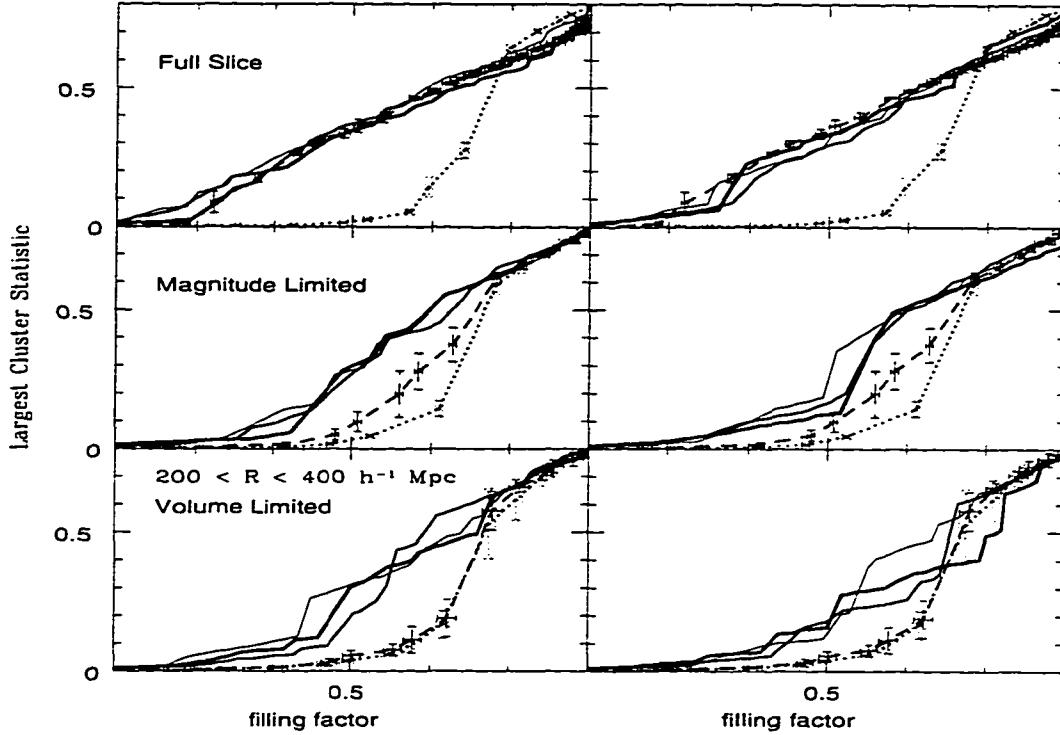


Figure 4.3: The Largest Cluster Statistic (LCS) results for the LCRS slices. The top panels are the results for the full slices ( $R \leq 600$  Mpc), the middle and bottom panels are the results for the magnitude and volume limited samples respectively ( $200 \leq R \leq 400$   $h^{-1}$  Mpc). The solid lines are the survey results (1 lightest, 3 darkest, for both north and south slices). The dashed lines are the LCS for corresponding Poisson Distributions corrected for selection and projection effects while the dotted lines are uncorrected Poisson Distribution results. The error bars are one sigma deviations over four realizations in all cases.

survey results. An explanation for this is that the distortion caused by the selection function produces a statistically inhomogeneous distribution. By separating the survey into two regions ( $60 \leq R \leq 400$   $h^{-1}$  Mpc and  $400 \leq R \leq 600$   $h^{-1}$  Mpc) the effect of the selection function is minimized and the LCS of the Poisson distribution shifts to the right. The reason is that in a magnitude limited survey the selection function is dependent on the distance from the observer and in the case of the LCRS peaks at a value of  $R \approx 200$   $h^{1-}$  Mpc. The middle panels show the results of a magnitude limited sample in the region  $60 \leq R \leq 400$   $h^{-1}$  Mpc.

There is now a clear distinction between survey and corrected Poisson distribution results. Results for the region  $400 \leq R \leq 600 h^{-1} \text{ Mpc}$  (not shown) are similar to those for the inner region shown. The slices percolate at lower filling factors than the standard in all cases. Both the outer and inner regions of the survey reveal a clear and unambiguous signal for a connected topology when analyzed separately.

The bottom panels show the results for volume limited subsamples (see Figure 4.4) derived from the survey in the region  $200 \leq R \leq 400 h^{-1} \text{ Mpc}$ .<sup>2</sup> Once again there is a clear signal at the one sigma level for a connected topology. The results above illustrate the necessity for volume limited samples. In order to compare percolation results they must come from statistically uniform samples. This conclusion contrasts with the method of de Lapparent, Geller and Huchra (1991) to remove the distortion effects of the selection function. In their work with the CfA slices they increase the size of outer lying grid cells in order to maintain a similar galaxy density throughout the survey. In essence, they may have traded one distortion for another: certainly the effects of their method are not understood and a comparison with their results is difficult. Sources of distortion in this study are statistically inhomogeneous distributions due to the selection function, two-dimensional analysis of three-dimensional surveys, the curvature of the slices and possible inherent galaxy incompleteness in the survey design. We can compensate for the selection effects by analyzing volume limited surveys and the projection effects by generating random two-dimensional reference catalogs that are similar to the surveys. Even then, the volume limited northern slices show stronger connectivity than the southern slices when the largest cluster statistic is examined over its entire range due either to curvature effects or regional variation in the topology at these scales. A three dimensional study will determine the cause of this difference. The results of this study imply a connected topology for all

---

<sup>2</sup>Volume limited samples with the same geometry as the magnitude limited samples were also analyzed. The results for those samples were similar to the results shown and the conclusions are the same.

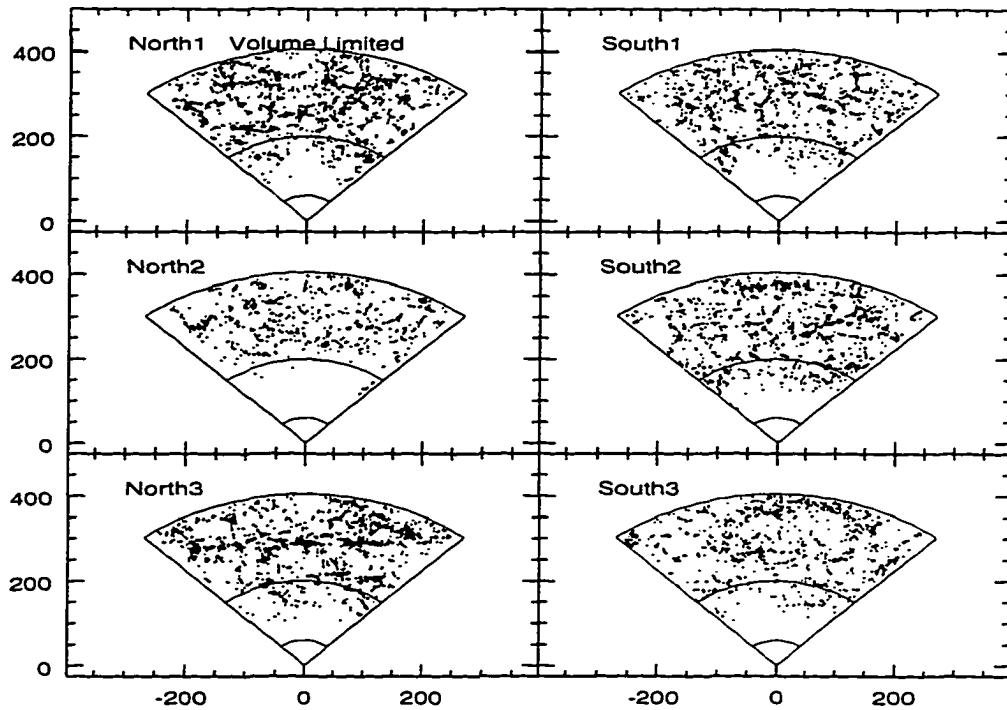


Figure 4.4: Maps of the volume limited distributions for all Las Campanas Redshift Survey slices ( $200 \text{ Mpc} \leq R \leq 400 \text{ } h^{-1} \text{ Mpc}$ ). The axis units are Mpc.

slices consistent with a filamentary geometry. A recent two-dimensional analysis of the LCRS found that the smoothed density distribution has a genus consistent with a random-phase Gaussian distribution of initial density fluctuations (Colley 1997). It has previously been shown that (Dominik & Shandarin 1992) the percolation results from highly smoothed, nonlinear, density distributions obtained from N-body simulations are consistent with Gaussian fields. Thus, the scenario arising from topological analysis of the Las Campanas galaxy survey is one of connected structures evolving from initial Gaussian perturbations of the density field consistent with inflationary cosmological models.

In Figure 4.4, the volume limited distributions for each slice are pictured. It is interesting to note that the largest cluster statistic is not predetermined by the initial filling factor for a non-Poisson distribution. The middle northern slice,

N2, is obviously the sparsest of the six slices (see Table 1) and yet it has the same degree of connectivity as the other slices. Our experience is that visual inspection is a poor method by which to judge the degree of connectivity or type of structures in a distribution. There is also the important result that the northern slices demonstrate stronger connectivity than the southern slices through percolation though this is hard to detect by visual inspection. Another important aspect of the Las Campanas survey is that there are small radial gaps in some of the slices. These radial voids are a cause a systematic error in the analysis and act to shift the largest cluster statistic to the right. The percolation analysis under these circumstances gives a lower estimate of the degree of connectivity.

Even though percolation analysis gives a clear prediction of the three dimensional topology, further examination of this data is necessary. We intend to expand our study to three dimensions with a complete analysis of uncertainties. In order to quantify the degree of connectivity in the survey, an assessment of edge and grid orientation effects must be made. Landy et al. have reported an enhancement of the power spectrum on length scales of roughly  $100 h^{-1}$  Mpc. This signal is associated with identifiable structures in the survey and is highly directional. Any effects on percolation results due to the directionality of the k-space fluctuations needs to be assessed by rotating the grid. Also, the consequences of possible redshift distortions must be evaluated and incorporated into the results. Fingers of God distortions are well known to astronomers but recent work by Praton and Melott (1997) has examined distortions that result in linear structures perpendicular to the line of sight. A comparison between the percolation of simulated galaxy catalogs in real space and their redshift space counterparts will help address this problem. In addition, the resolution of the LCRS must be optimized in terms of grid cell size and uncertainties in the redshift positions. We intend to report on these studies in a forthcoming paper.

# Chapter 5

## Conclusion

### 5.1 Summary of Results

This thesis demonstrates the potential of percolation analysis as a statistical tool in the field of cosmology. We have shown it capable of assessing the topology and geometry of evolved distributions. We have also demonstrated that over a wide and realistic range of spectral index of the power spectrum the overdense regions in N-Body simulations reveal topologies that correspond to a network or filamentary structure. In addition, simulations described as having relatively more power on large scales ( $n \leq -1$ ) show a definite trend towards a bubble topology. It is important to note that these studies were done in real space, but the conclusions drawn from the study hold generally. The results are not meant to be compared with results from survey studies in redshift space but were used to calibrate and test the parameters of percolation. The comparison of theoretical models with data requires and accurate modeling of galaxy distributions and is beyond the scope of this thesis. We also demonstrated that the number of clusters statistic is sensitive to the non-Gaussianity and slope of the power spectrum of a mass distribution. This property can be exploited for fields with sufficient resolution to estimate the slope of the power spectrum in mass distributions.

The direct application of percolation analysis to mass and pointwise distributions derived from astronomical surveys has produced results that are significant and complimentary to other studies. The IRAS 1.2 Jy Redshift Survey as represented by a Wiener reconstruction of the mass density is shown to have a well connected void distribution and a somewhat clumpy mass distribution in the local vicinity. This result must be evaluated with the understanding that the survey is small (in statistical terms) and represents a sparse sample of the local galaxy distribution. In contrast, the regions surveyed by the Las Campanas Redshift Survey display a galaxy distribution consistent with a network topology. This is the first confirmation of a network structure to such a depth ( $R = 400 h^{-1}$  Mpc) and over such broad and separated regions of space. These studies also revealed the shortcomings of percolation analysis. The methods developed to apply percolation to cosmological applications require significant resolution for density fields and number density of galaxies for pointwise distributions. For galaxy distributions the inherent selection function associated with surveys introduces a statistical inhomogeneity that requires volume limited subsampling to remove. The initial number density of the survey must be high enough to withstand this reduction in information and still provide a signal that is above the level of random noise. The next phase in the development of percolation is to systematically study these shortcomings and other errors in the analysis in order to quantify the uncertainties of the results.

In short, this study adds evidence to the prevailing view that the large scale structure of the local universe developed from initially small perturbations in the matter density field at the time of recombination. Those perturbations grew by gravitational attraction to produce a connected structure of galaxies in regions spanning hundreds of megaparsecs. The conclusions are also support scenarios where there is significant power on large scales ( $1-100 h^{-1}$  Mpc) in the primordial power spectrum.

## **5.2 Error Analysis and Percolation**

To date, the conclusions of percolation analysis have relied on distinctions in the parameters between test distributions and statistically random distributions. The errors have been approximated by standard deviations of independent random samples. These errors reflect the statistical variation of random distributions, but errors in the survey distributions also need to be addressed. These errors include uncertainties in the redshift positions of the galaxies and redshift distortions. The redshift distortions inherent in galaxy surveys raise important questions about the relevance of the conclusions of the LCRS study. It is important to realize that there will be refinements in astronomical surveys which in the future which may include corrections for redshift distortions. It is also important to analyze the data that exists presently in order to help suggest improvements in future survey designs and test the sensitivity and applicability of the percolation parameters. There are also systematic errors associated with the methodology of percolation that need to be assessed. The effects, if any, of grid cell orientation and size need to be examined. There are certainly boundary effects integral to the percolation method because no account is made of information outside the survey boundary. Assessments of all errors applicable to percolation analysis will clarify the conclusions of the process and add to the qualitative aspect of the conclusions.

## **5.3 Future Applications**

Percolation analysis has already provided valuable information in the area of large scale structure. It is fortunate that the shortcomings of percolation analysis are due to shortcomings in astronomical surveys. As astronomers increase the speed of gathering data, surveys larger in volume and galaxy numbers will mitigate the statistical errors in the analysis. An increase in survey volume will also help insure the statistical fairness of the results. These expectations will be checked soon by

the analysis of the upcoming Sloan Digital Sky Survey.

# References

- Arnold, V. I., Shandarin, S. F., & Zel'dovich, Ya. B. 1982. *Geophys. Astrophys. Fluid Dyn.*, 20, 111
- Babul, A., & Starkman, G. D. 1992 *ApJ*, 401, 28
- Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986. *ApJ*, 304, 15
- Baugh, C. 1993. *MNRAS*, 264, 87
- Bhavsar, S., & Barrow, J. 1983. *MNRAS*, 205, 61
- Bhavsar, S. P., & Splinter, R. J. 1996 *MNRAS*, 282, 1461
- Bouchet, F. 1995, talk at the 15-th Moriond Astrophysics Meeting on *Clustering in the Universe*
- Bunn, E., Fisher, K. B., Hoffman, Y., Lahav, O., Silk, J., & Zaroubi, S. 1994. *ApJ Lett.*, 432, L75
- Coles, P. & Lucchin, F. 1995 *Cosmology: The Origin and Evolution of Cosmic Structure* Wiley & Sons (New York).
- Coles, P., Melott, A. L., & Shandarin, S. F. 1993. *MNRAS*, 260, 572
- Colley, W. N. 1997, *astro-ph/9612106*
- Davis, M., Efstathiou, G., Frenk, C., & White, S. D. M. 1985. *ApJ*, 292, 371
- Davis, M., Huchra, J., Latham, D. W., & Tonry, J. 1982. *ApJ*, 253, 423
- Dekel, A., & West, M. J. 1985, *ApJ*, 288, 411
- Dominik, K., & Shandarin, S. F. 1992, *ApJ*, 393, 450
- Doroshkevich, A. G. 1970, *Astrophysics*, 6, 320
- Einasto, J., Klypin, A. A., Saar, E., & Shandarin, S. F. 1984. *MNRAS*, 206, 529
- Fisher, K. B., Huchra, J. P., Strauss, M. A., Davis, M., Yahil, A., & Schlegel, D. 1995b, *ApJ Suppl.*, 100, 69
- Fisher, K. B., Lahav, O., Hoffman, Y., Lynden-Bell, D., & Zaroubi, S. 1995a *MNRAS*, 272, 885

- Ganga, K., Cheng, E., Meyer, S., & Page, L. 1993 *ApJ Lett.* 410, L57
- Geller, M. J. & Huchra, J. 1989, *Science*, 246, 897.
- Giovanelli, R., & Haynes, M. P. 1985, *AJ*, 90, 2445
- Gott, J. R., Melott, A. L., & Dickinson, M. 1986, *ApJ*, 306, 341
- Gott, J. R., Miller, J., Thuan, T. X., Schneider, S. E., Weinberg, D. H., Gammie, C., Polk, K., Vogeley, M., Jeffery, S., Bhavsar, S., Melott, A. L., Giovanelli, R., Haynes, M. P., Tuily, R. & Hamiltion, A. 1989, *ApJ*, 340, 625
- Gunn, J. E. & Weinberg, D. H. 1995, *Wide-Field Spectroscopy and the Distant Universe*, eds. S. J. Maddox and A. Aragón – Salamanca. (Singapore : World Scientific).
- IRAS* Science Working Group (Washington D.C.: U.S. Government Printing Office) (PSC)
- Hamilton, A. J. S., Gott J. R.. & Weinberg D. 1986, *ApJ*, 309, 1
- Kirshner, R. P., Oemler, A., Schechter, P. L., & Shectman, S. A. 1981, *Apj*, 248, L57
- Klypin, A. A. 1987, *Soviet Astron.*, 31, 8
- Klypin, A. A., & Shandarin, S. F. 1983, *MNRAS*, 20, 891
- Klypin, A. A., & Shandarin, S. F. 1993, *ApJ*, 413, 48
- Kofman, L., Pogosyan, D., Shandarin, S. F., & Melott, A. L. 1992, *ApJ*, 393, 437
- Kolb, E. W. & Turner, M. S. 1990 *The Early Universe* Addison-Wesley (New York).
- Landy, S. D., Shectman, S. A., Lin, H., Kirshner, R. P., Oemler, A. A., & Tucker, D. 1996 *ApJ Lett.*, 456, 1
- Lahav, O., Fisher, K. B., Hoffman, Y., Schaarf, C. A., & Zaroubi, S. 1994 *ApJ Lett.*, 423, L93
- de Lapparent, V., Geller, M., & Huchra, J. P. 1986, *ApJ*, 301, L1
- de Lapparent, V., Geller, M., & Huchra, J. P. 1991, *ApJ*, 369, 273
- Lawrence, A., Rowan-Robinson, M., Crawford, J., Saunders, W., Ellis, R. S..

- Frenk, C., Parry, I., Efstathiou, G., & Kaiser, N. 1995 in preparation
- Lin, H., Kirshner, R. P., Shectman, S. A., Landy, S. D., Oemler, A., Tucher, D. L., & Schechter, P. L. 1996a, ApJ, 471, 617
- Linde, A. D. 1990, *Particle Physics and Inflationary Cosmology* (New York : Harwood).
- Luo, S., & Vishniac, E. 1995, ApJ Suppl., 96, 429
- Matsubara, T. 1994, ApJ, 434, L43
- Mecke, K., Buchert, T., & Wagner, H. 1994, A&A, 288, 697
- Melott, A. L. 1986, Phys. Rev. Lett., 56, 1992
- Melott, A. L. 1990, Physics Reports, 193, 1
- Melott, A. L., & Shandarin, S. F. 1993, ApJ, 410, 469
- Melott, A. L., Weinberg, D. H., & Gott J. R. 1988, ApJ, 328, 50
- Mo, H. J., & Börner, G. 1990, A&A, 238, 3
- Moore, B., Frenk, C. S., Weinberg, D. H., Saunders, W., Lawrence, A., Ellis, R. S., Kaiser, N., Efstathiou, G., & Rowan-Robinson, M. 1992, MNRAS, 256, 477
- Pauls, J. L., & Melott, A. L. 1995, MNRAS, 274, 99
- Padmanabhan, T. 1993 *Structure Formation in the Universe* Cambridge University Press (Cambridge, England).
- Peebles, P. J. E. 1980 *The Large Scale Structure of the Universe* Princeton University Press (Princeton, NJ).
- Praton, E. A., Melott, A. L., & McKee, M. Q. 1997, ApJ Lett., 479, L15
- Protogeros, Z. A. M., & Weinberg, D. H. 1997, astro-ph/9701147
- Ryden, B. S. 1988, ApJ Lett., 333, L41
- Rybicki, G. B., & Press, W. H. 1992, ApJ, 398, 169
- Sahni, V., Sathyaprakash, B. S., & Shandarin, S. F. 1994, ApJ, 431, 20
- Sahni, V., Sathyaprakash, B. S., & Shandarin, S. F. 1997, ApJ Lett., 476, L1
- Sathyaprakash, B. S., Sahni, V., & Shandarin, S. F. 1996, ApJ Lett., 462, L5
- Sathyaprakash, B., Sahni, V., & Shandarin, S. F. 1997, in preparation

- Shandarin, S. F., Melott, A. L., McDavitt, K., Pauls, J. L., & Tinker, J. 1995.  
*Phy. Rev. Lett.*.. 75, 7
- Shandarin, S. F. 1983, Soviet Astron. Lett., 9, 104
- Shandarin, S. F. 1994, in the Proceedings of the Workshop on "Large-Scale Structure in the Universe: Theoretical and Observational Aspects". ed. J. Mücke, in press
- Shandarin, S. F. 1995, in the Proceedings of the Fifteenth Moriond Astrophysics Meeting on *Clustering in the Universe*, ed. C. Balkowski, S. Maurogordato and J. Trần Thanh Vân (Editions Frontières), in press
- Shandarin, S. F., & Zel'dovich, Ya. B. 1983. *Comments Astrophys.*.. 10 33
- Shane, C. D. & Wirtanen, C. A., 1967. *Publ. Lick Obs.*.. 22, 1.
- Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., Kirshner, R. P., Lin, H.. & Schechter, P. L. 1996, *Wide-Field Spectroscopy and the Distant Universe. proceedings of the 35th Herstmonceux Conference*. eds. S. J. Maddox & A Aragón-Salamanca. World Scientific. Singapore
- Smoot, G. F., Bennett, C. L., Kogut, A., Wright, E. L., Aymon, J., Boggess, N. W., Cheng, E. S., De Amici, G., Gulkis, S., Hauser, M. G., Hinshaw, G., Jackson, P. D., Janssen, M., Kaita, E., Kelsall, T., Keegstra, P., Lineweaver, C., Loewenstein, K., Lubin, P., Mather, J., Meyer, S. S., Moseley, S. H., Murdock, T., Rokke, L., Silverberg, R. F., Tenorio, L., Weiss, R., & Wilkinson, D. T. 1992 *ApJ Lett.*, 396, L1
- Stauffer, D. & Aharony, A. 1992 *Introduction to Percolation Theory* Taylor & Francis (London, Washington , DC).
- Strauss, M. A., Huchra, J. P., Davis, M., Yahil, A., Fisher, K., & Tonry, J. 1992a.  
*ApJ Suppl.*, 83, 29
- Tomita, H. 1986, *Prog. Theor. Phys.*, 76, 952
- Tully, R. B., & Fisher, J. R. 1987, *Nearby Galaxies Atlas* (Cambridge: Cambridge University Press).

- Vishniac, E. 1986, in Inner Space/Outer Space eds. E. W. Kolb et al. (Chicago: University of Chicago Press), 1990
- Vogeley, M. S., Park, C., Geller, M. J., Huchra, J. P., & Gott, J.R. 1994. ApJ. 420, 525
- Weinberg, D. H., Gott, J. R., & Melott, A. L. 1987, ApJ. 321, 2
- Yess, C., & Shandarin S. F. 1996, ApJ, 465, 2
- Yess, C., Shandarin, S. F., & Fisher K. B. 1997. ApJ, 474, 553
- Zaroubi, S., Hoffmean, Y., Fisher, K. B., & Lahav, O. 1995. ApJ, 449, 446
- Zel'dovich, Ya. B. 1982, Soviet Astron. Lett., 8, 102
- Zwicky, F., Herzog, E., Wild, P., Karpowicz, M. & Kowai, C. T.. 1961-1968. *Catalogue of Galaxies and Clusters of Galaxies*. California Institute of Technology. Pasadena