# ABSTRACT

Title of dissertation: EVOLUTION OF DISCRETE DYNAMICAL SYSTEMS

Jonathan Ozik, Doctor of Philosophy, 2005

Dissertation directed by: Professor Edward Ott
Department of Physics

We investigate the evolution of three different types of discrete dynamical systems. In each case simple local rules are shown to yield interesting collective global behavior.

(a) We introduce a mechanism for the evolution of growing small world networks. We demonstrate that purely local connection rules, when coupled with network growth, can result in short path lengths for the network as a whole.

(b) We consider the general character of the spatial distributions of populations that grow through reproduction and subsequent local resettlement of new population members. Several simple one and two-dimensional point placement models are presented to illustrate possible generic behavior of these distributions. We show, both numerically and analytically, that all of the models lead to multifractal spatial distributions of population.

(c) We present a discrete lattice model to investigate the segregation of three species granular mixtures in horizontally rotating cylinders. We demonstrate that the simple local rules of the model are able to reproduce many of the experimentally observed global phenomena.

# EVOLUTION OF DISCRETE DYNAMICAL SYSTEMS

by

Jonathan Ozik

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2005

Advisory Commmittee:

Professor Edward Ott, Chair/Advisor
Professor Brian Hunt
Professor Wolfgang Losert
Professor Jim Yorke
Professor Michael Coplan

# DEDICATION

To my parents, my sister, and the cats

# ACKNOWLEDGMENTS

There are many people that deserve my gratitude for making my journey through graduate school both possible and enjoyable.

First of all, I would like to thank my parents for the love and support that they have given me over the years. None of this would have been possible without them. To Dana, the best little sister a big brother could have, I would like to say thank you for being such an inspiring model of hard work and perseverance.

A deep, heartfelt thanks goes to Sarah Newport for not only putting up with me but also for providing the indispensable constant encouragement, as well as for being the wonderful person that she is.

I was extremely fortunate to have had Professor Edward Ott as my advisor. His willingness to take on interesting problems in a variety of areas combined with his sharp insights allowed for a research experience that was both wide in scope and intellectually rewarding. I am forever grateful for his guidance. I would also like to thank Professor Brian Hunt and Professor Wolfgang Losert for the fruitful collaborations and the invaluable experiences gained through my interactions with them.

I was very lucky to have met so many wonderful people in my time at Maryland. They made the various difficulties that one comes across while pursuing a Ph.D. more bearable. Thanks goes to David Noyes for being such a good friend and for listening to my many rants. I would also like to thank Dušan Turčan for being an integral part, from start to finish, of my experience here. Andrew Berkley and Tobias Dŭrkop deserve the credit for keeping me sane during my first year here. Nobumi Katsu, Boris Gelman, Safraz Ishmael, Adrienne Anderson, Ellen O'Connell, Myra Flitcroft, Andy Tillotson, Jennifer Hains, Brendan Foster, Tim McCaskey, Ellen Roche were some of the people who, at different periods of my graduate school career, made my time outside of work

very entertaining.

I would like to acknowledge my A.V. Williams Chaos group colleagues who contributed to the friendly work environment: Seung-Jong Baek, Juan Restrepo, Henry Zheng, Yue-Kin Tsang, Romulus Breban, Doug Armstead. Thanks also goes to the administrative staff of IREAP for all of their hard work.

The staff in the Physics Department were some of the nicest and most helpful people that I met at Maryland. Thanks go to Jane Hessing, Lorraine DeSalvo, Bernie Kozlowski, the people in the Chair's Office, and the people in the Payroll and Business offices.

Finally, a large debt of gratitude goes to Giogio and Jiji, my loving cats, who always helped me put seemingly daunting tasks into perspective.

TABLE OF CONTENTS

List of Tables vii

List of Figures viii

1 Introduction 1

2 Growing networks with geographical attachment preference: Emergence of small worlds 3

2.1 Introduction . . . 3

2.2 Growing Network Model . . . 5

2.3 Degree Distribution . . . 5

2.4 Clustering . . . 8

2.5 Characteristic Path Length . . . 10

2.6 Geography . . . 12

2.7 Conclusion . . . 16

3 Formation of Multifractal Population Patterns from Reproductive Growth and Local Resettlement 17

3.1 Introduction . . . 17

3.2 Models and Results . . . 21

3.2.1 1-D Random Interval (Model 1) . . . 21

3.2.2 2-D Random Square (Model 2) . . . 24

3.2.3 1-D Larger Interval (Model 3) . . . 28

3.2.4 2-D Sparse Square (Model 4) . . . 28

3.2.5 2-D Unstructured (Model 5) . . . 31

3.2.6 2-D Square (Model 6) . . . 32

3.2.7 2-D Triangle (Model 7) . . . 33

3.3 The Effect of Clipped Data on Multifractality . . . 37

3.4 Inhomogeneity . . . 40

3.5 Theory . . . 43
3.5.1 Model 1 . . . 43
3.5.2 Model 2 . . . 44
3.5.3 Models 6 and 7 . . . 47
3.6 Conclusion . . . 48

4 Simulating the Segregation of Ternary Granular Mixtures in Long Rotating Cylinders 49
4.1 Introduction . . . 49
4.2 The Model . . . 52
4.3 Simulation Results and Discussion . . . 55
4.4 Conclusion . . . 60

A Calculation of $\bar{H}(k)$ 62

B Behavior of $\langle l|k\rangle$ 64

C Calculation of Model Dimension Spectra 65
C.1 The Partition Function Formalism . . . 65
C.2 Application to the Models . . . 67

D Proof of Theorem from Appendix C 69

Bibliography 70

LIST OF TABLES

3.1 A summary of numerical and theoretical results for the point placement models and the Earth at Night image (Fig. 3.1). . . . . 35

3.2 Summary of reproduction and local resettlement strategies employed in Models 1-7. 36

LIST OF FIGURES

2.1 Our growing network model, illustrated for $m = 2$. We begin with $m+1$ completely connected nodes on the circumference of a circle (top left). At each subsequent time step we: a) add a new node in a randomly chosen inter-node interval along the circle circumference, with every interval having equal probability of being chosen, and b) connect the new node to its $m$ nearest neighbors, with nearest here referring to distance along the circle circumference. Steps (a) and (b) are repeated until the desired system size is reached. . . . . 6

2.2 The open circles represent the degree distribution $P(k)$ for a network grown according to our model with $N = 10^5$ and the solid line is the analytically calculated ensemble averaged degree distribution [Eq. (2.2)], both with $m = 2$. . . . . 7

2.3 Average clustering, $\langle C \rangle$, versus system size, $N$, for a simulated network with $m = 2$. As $N$ grows, the average clustering approaches the value (dashed line) predicted in Eq. (2.4). . . . . 9

2.4 Semilogarithmic graph of the characteristic path length, $L$, versus the system size, $N$. The data show the small world slow path length growth characteristic, $L \sim \ln N$. The straight line is a fit to the data. . . . . 10

2.5 An illustration of network growth in our model for $m = 2$. The network starts off (left) with three adjacent nodes, labeled by $A$, $B$, and $C$, connected to each other via links. When the network reaches a network size of 100 nodes (right), the original three labeled nodes are no longer adjacent, but have been "pushed apart" by the new nodes that were inserted between them. The links connecting $A$, $B$, and $C$ serve as short cuts (similar to the shortcuts in the Watts-Strogatz model [4]), resulting in a small characteristic path length for the network. . . . . 11

2.6 An illustration of the different regions associated with node $B$. Let node $B$ have degree $k$ and longest CCW link $l$. A new node inserted in region IV will not affect either $k$ or $l$. A new node inserted in region III (between the CCW nearest neighbor $C$ and the CCW farthest connected node $D$) increases $l$ by one but leaves $k$ unchanged. A new node inserted in region II (between nodes $B$ and $C$) increases both $k$ and $l$ by one. Since region I (between nodes $A$ and $B$) is in the clockwise direction, a new node inserted here increases $k$ by one but leaves $l$ unchanged. . . . 13

3.1 A version of the Earth at Night image [3]. This is a composite image created with data from the Defense Meteorological Satellite Program of the Earth's nighttime man-made lights as seen from space. For the analysis, an 8-bit grayscale 2400x800 pixel image was used, where light from the reflective landmasses found in the original image was removed with an appropriate uniform background subtraction. The image pixels have intensity values ranging from 0 to 218. . . . . 18

3.2 Plot for calculating the information dimension $D_1$ [see Eq. (3.3)] for the Earth at Night image. The quantity $\sum_{i=1}^{N(\epsilon)} \mu_i \ln \mu_i$ is plotted versus $\ln \epsilon$ (open circles), where $\epsilon$ is in units of pixels, resulting in a linear *scaling region*, the slope of which is a finite-scale approximation of $D_1$. The solid line is a linear fit to the data with a slope of $1.60 \pm 0.02$. . . . . 19

3.3 Three representative steps of Models 1 and 3. . . . . 22

3.4 Semilog plots of histograms of point locations on the unit interval for a $10^5$ point distribution generated by Model 1, for various bin sizes $\Delta x$. a) $\Delta x = 1/16$, b) $\Delta x = 1/32$, c) $\Delta x = 1/64$, d) $\Delta x = 1/128$. . . . . 23

3.5 Plot of $D_q$ for Model 1. The open circles are the numerical values calculated from the slope of a line fit to $(1-q)^{-1}\ln I(q,\epsilon)$ versus $\ln(1/\epsilon)$ [see Eq. (3.1)], averaged over 20 realizations of $10^6$ point distributions (except for $q=1$, in which case the line is fit to $\sum_{i=1}^{N(\epsilon)}\mu_i\ln\mu_i$ versus $\ln\epsilon$ [see Eq. (3.3)]), with the error bars indicating the sample standard deviation for each value. The solid line is the analytical result of Eq. (3.14). As an example the inset shows the determination of $D_1$ for one point distribution. . . . . 25

3.6 Three representative steps of Models 2 and 4. . . . . 26

3.7 A plot of a $4\times10^5$ point distribution on the unit square generated by Model 2. Dense regions are magnified to illustrate the heterogeneity of the point densities in the distribution at different scales. . . . . 27

3.8 Plot of $D_q$ for Model 2. The open circles are the average numerical values of $D_q$ (see caption in Fig. 3.5) obtained from 20 realizations of $10^6$ point distributions generated by this model and the error bars are the sample standard deviations for each value. The solid line is the analytical result of Eq. (3.28). . . . . 27

3.9 Semilog plots of histograms of point locations on the unit interval for a $10^5$ point distribution generated by Model 3, for various bin sizes $\Delta x$. a) $\Delta x = 1/16$, b) $\Delta x = 1/32$, c) $\Delta x = 1/64$, d) $\Delta x = 1/128$. . . . . 29

3.10 Illustration of type I (labeled I) and type II (labeled II) squares, where black dots indicate populated vertices. . . . . 30

3.11 A plot of a $10^5$ point distribution on the unit square generated by Model 4. . . . . 31

3.12 A plot of a $10^5$ point distribution on the unit square generated by Model 5. . . . 32

3.13 Illustration of Step (b) of Model 6. . . . . 33

3.14 Plot of $D_q$ for Model 6. The open circles are the average numerical values of $D_q$ (see caption in Fig. 3.5) obtained from 20 realizations of $10^6$ point distributions generated by this model and the error bars are the sample standard deviations for each value. The solid line is the analytical result of Eq. (3.34). . . . . 34

3.15 Illustration of Step (b) of Model 7. . . . . 35

3.16 a) Plot of $D_q$ for the Earth at Night image. b) Plot of $D_q$ for a $10^6$ point distribution generated by Model 5 before the clipping procedure discussed in Section 3.3 is applied, and c) after it is applied (note the different vertical scales). The error bars reflect the uncertainty involved in determining the slopes of the scaling regions in the plots of the quantity $(1-q)^{-1}\ln I(q,\epsilon)$ versus $\ln(1/\epsilon)$ [see Eq. (3.1)]. . . . . 38

3.17 Histogram of individual pixel intensities for the Earth at Night image. . . . . 39

3.18 Histograms of the number of points in individual simulated pixels (boxes in an $\epsilon = 1/1400$ grid) used to cover a $10^6$ point distribution generated by Model 5, a) for the original distribution and b) after the original distribution was clipped by allowing a maximum of 58 points in any simulated pixel. . . . . 41

3.19 A plot of the Weierstrass function $f_w(z)$ [Eq. (3.6)]. In our numerical implementation the infinite upper limit on the summation in Eq. (3.6) is replaced by $j_{max}$, where $j_{max}$ is chosen such that the smallest value of $\epsilon$ used in determining $D_q$ is at least one order of magnitude larger than the smallest scale of the roughness, $3^{-j_{max}}$. 42

4.1 The S-shaped profile of a flowing granular mixture in a rotating cylinder. . . . . 50

4.2 Picture of the "bands within bands" structures observed by Newey et al. [43]. The light blue particles are the largest, the dark blue particles are the intermediate sized ones and the green particles are the smallest. . . . . 52

4.3 Schematic diagram of the discrete rotation process. The lattice is separated into its back ($x < N_x/2$) and front ($x \geq N_x/2$) halves. The back half is raised one step and the bottom slice of the front half is rotated and attached to the bottom of the back half. The front half is then lowered one step. . . . . 53

4.4 An illustration of the adjacent lattice sites, represented by gray cubes, which are included in the summation for calculating the generalized friction $F$ (see Eq. (4.2)) experienced by a particle (white cube). . . . . 55

4.5 a) View of the three dimensional lattice after $2 \times 10^5$ time steps. Initially the lattice is loaded with a uniformly mixed state of particles A, B and C (colored red, yellow and black, respectively), with equal proportions. b) Profiles of the particle mixture surface at three different axial ($y$) positions. [$F_{AA} = 0.4$, $F_{AB} = 0.44$, $F_{BB} = 1.0$, $F_{AC} = 0.9$, $F_{BC} = 1.5$, $F_{CC} = 2.4$, $R = 2.5 \times 10^4$] . . . . . 57

4.6 a) View from the rear ($x = 0$) of the lattice in Fig. 4.5(a), showing the "bands within bands" structures. This perspective most closely resembles the laboratory view of the surface of a rotating granular mixture (cf. Fig. 4.2). b) The relative abundance $\phi$ of each of the particle types for two dimensional slices ($0 \leq x < N_x$ and $0 \leq z < 2N_z$) at each axial position $y$. [Black signifies particles A, yellow signifies particles B and red signifies particles C.] . . . . . 58

4.7 a) Spacetime plot of the our simulation for a total of $5 \times 10^5$ timesteps. In each time step, the most abundant particle type in the rear ($x = 0$) slice of the lattice is picked for each axial ($y$) position. The black, yellow, and red regions correspond to particles A, B, and C respectively. The relative abundances of the particles are 25% A particles, 30% B particles and 45% C particles. b)-e) Top views of the lattice corresponding to different times in the space-time plot. [(b) $t = 0$ (c) $t = 2 \times 10^4$, (d) $t = 2 \times 10^5$, (e) $t = 5 \times 10^5$] . . . . . 59

# Chapter 1

## Introduction

The study of complexity and complex systems has become increasingly popular in many scientific disciplines. Complex systems are collections of interacting subunits, where the global behavior of the whole is difficult to infer from the properties and interactions of the subunits. Thus computers, neural networks, human societies and weather are all examples of complex systems. An important realization and a unifying concept in the study of such systems is that complex global behavior does not require complicated interactions between the individual subunits. In fact, it has been repeatedly shown (e.g. cellular automata [1], agent based models [2]) that simple building blocks can result in very complicated, often surprising and, perhaps most importantly, potentially meaningful global behavior.

As a result, researchers in diverse fields such as sociology, computer science, biology and physics, just to name a few, are modeling complexity with simply interacting simple subunits. This, in turn, has led to a large amount of interdisciplinary research, where traditionally disconnected fields find common ground under this new framework.

Discrete time/discrete state models (DTDSMs) can be very useful in describing a variety of complex systems. A DTDSM is any model that has a discrete state space and which is updated at discrete time steps. The prototypical examples of DTDSMs are cellular automata; however, there are many other types of models that can be classified as DTDSMs (e.g., incrementally growing models, discrete lattice models, etc.). One of the aspects that makes these models appealing is the simplicity with which they can be simulated on a computer. Numerical solutions of complicated continuum models will almost always involve some form of discretization procedures and, thus, are approximate solutions to the specified continuous models. DTDSMs, on the other hand, can be simulated exactly even for very intricate models. Additionally, since, as noted above, complexity can emerge from a collection of simply interacting entities, the DTDSM describing a complex system can be rather simple, sometimes even allowing for analytical approaches. In this dissertation we present several DTDSMs applied to a number of complex systems.

In Chapter 2 we introduce a simple mechanism for the evolution of small world networks. Our model is a growing network in which all connections are made locally to geographically nearby sites. Although connections are made purely locally, network growth leads to stretching of old connections and to high clustering. Our results suggest that the abundance of small world networks in geographically constrained systems is a natural consequence of system growth and local interactions.

In Chapter 3 we consider the general character of the spatial distribution of a population that grows through reproduction and subsequent local resettlement of new population members. Several simple one and two-dimensional point placement models are presented to illustrate possible generic behavior of these distributions. It is shown, numerically and analytically, that these models all lead to multifractal spatial distributions of population. Additionally, we make qualitative links between our models and the example of the Earth at Night image [3], showing the Earth's nighttime man-made lights as seen from space. The Earth at Night data suffer from saturation of the sensing photodetectors at high brightness ('clipping'), and we account for how this influences the determined multifractal dimension spectrum of the light intensity distribution.

In Chapter 4 we investigate the segregation of three species granular mixtures in horizontally rotating cylinders. We present a discrete lattice model based on the different frictional properties of the three particle types and show that it yields results that are qualitatively similar to those of the physical experiment.

Chapter 2

Growing networks with geographical attachment preference: Emergence of small worlds

## 2.1 Introduction

Recently there has been considerable interest in the classification of physical systems according to the topological properties of the networks to which they map (e.g., [4–13]), where the constituent parts are modeled as nodes and links between nodes denote some type of interaction (for reviews see [14–17]). This method of classification has the potential to shed light on underlying organizational principles. In this spirit, we focus on the "small world" network topology introduced by Watts and Strogatz [4].

Here we represent a network as an undirected graph: a collection of $N$ points (*nodes*) with connections (*links*) between some pairs of them. If two nodes are connected, we say that they are *neighbors*. We call the number of connections to node $i$ the *degree* of node $i$ and we denote it $k_i$.

Small world networks are characterized by two main properties. First, their characteristic path length, $L$, grows as $\ln N$ or slower, similar to an Erdős-Rényi (ER) random network. The characteristic path length is the smallest number of links connecting a pair of nodes, averaged over all pairs of nodes. Second, the network has a high average clustering compared to an ER random network of equal size and average node degree. The clustering $C_i$ of node $i$ is defined by $C_i = q_i/[(1/2)k_i(k_i - 1)]$, where $q_i$ is the total number of links between the $k_i$ neighbors of node $i$, and $(1/2)k_i(k_i - 1)$ is the maximum number of links that could exist between $k_i$ nodes. Networks exhibiting small world characteristics are found in many varied fields of research. Some examples of such networks are the neuronal network of the worm *C. elegans*, the electric power grid of southern California, and the friendship network of Madison Junior High School students [5].

The Watts and Strogatz model is the following prescription for creating a small world network. The initial state has a fixed number of nodes equally spaced on the circumference of a circle. Each node is linked to its $m$ nearest neighbor nodes, where $m$ is even and nearest here refers to the distance along the circumference of the circle. In this way a regular network with a large average clustering is created. Next, a proportion $p$ of the links are chosen at random and

"rewired" such that one end of the link is kept fixed and the other end is linked to a randomly chosen node. These random links can serve as short cuts across the circle, drastically decreasing the characteristic path length of the network. It was found [4] that, for a relatively small rewiring probability $p$, the characteristic path length of the network becomes comparable to that of an ER random network, while the network still maintains a high average clustering.

The network construction of the Watts and Strogatz model very nicely illustrates the small world property, and, furthermore, it is probably a reasonable model for how some networks are formed. However, the small world property is, of course, much more general than their particular example, and it is useful to study other mechanisms for forming small world networks. In particular, we will be interested in networks that grow in time from small size to large size by the successive addition of new nodes (see [8, 14–17] for other models of growing networks).

Many networks have their topology influenced by geographical constraints. The nodes are separated by some physical distance and thus their ability to know the complete state of all the network nodes at a given time is restricted. Consequently, in our model we restrict the formation of links between nodes to result from geographically local processes. That is, when a new node appears, it forms links only to those preexisting nodes that are geographically close to it. In spite of the link formation being exclusively local, long-range links will be shown to arise as a result of network growth. This, in addition to the clustering induced by local connections, yields the small world property.

We say that a growing network model has the small world property if it satisfies the following three criteria as the number of nodes $N \to \infty$:

a) Small average node degree, $\langle k \rangle = O(1)$;

b) Small characteristic path length, $L \sim \ln N$;

c) High average clustering, $\langle C \rangle = O(1)$ (i.e., $\langle C \rangle$ does not decay with $N$).

Part (a) of the definition is included to ensure that highly connected networks that trivially satisfy criteria (b) and (c) are not considered to be small world networks (e.g., if every node is connected

to every other node then $L = \langle C \rangle = 1$ but $\langle k \rangle = N - 1$).

## 2.2 Growing Network Model

Our model, as mentioned above, is a growing network. We begin with an initial state of $m + 1$ all-to-all connected nodes on the circumference of a circle (Fig. 2.1). (We take $m$ to be even.) We note that this initial state is chosen solely for convenience and it has no effect on the long time network properties. At each subsequent discrete time step we grow the network according to the following prescription: a) a new node is placed in a randomly chosen inter-node interval along the circle circumference, where all intervals have the same probability of being chosen; b) the new node makes $m$ links to its $m$ (previously existing) nearest neighbors. Nearest here refers to the distance measured in number of intervals along the circumference of the circle.

These steps are repeated sequentially, creating a network with a temporally growing number of nodes $N$. We note that, since the network size, $N$, is incremented by one with each discrete time step, $N$ can be used interchangeably as a system size or a time variable.

## 2.3 Degree Distribution

We now calculate the degree distribution for our network when $N$ is large. We define $\hat{G}(k, N)$ as the number of nodes with degree $k$ when the system size (or time) is $N$. Since all new nodes are initially created with $k = m$, and links can only be added to nodes, $\hat{G}(k, N) = 0$ for $k < m$. At time $N$, a node with degree $k = m$ is added to the network, and if it links to a previously existing node $i$, then $k_i \to k_i + 1$. Each preexisting node is equally likely to be connected to the new node, and therefore the probability that a given preexisting node has its degree increased by 1 is $m/N$.

We now take the average over all realizations of the possible random placements of the new node. This yields the following evolution equation for the average of $\hat{G}$, which we denote $G$:

$$G(k, N+1) = \left(1 - \frac{m}{N}\right) G(k, N) + \frac{m}{N} G(k-1, N) + \delta_{km}, \tag{2.1}$$

where $\delta_{km}$ is the Kronecker delta function. The first term on the right hand side is the expected number of nodes with degree $k$ at time $N$ whose degrees remain the same at time $N + 1$. The

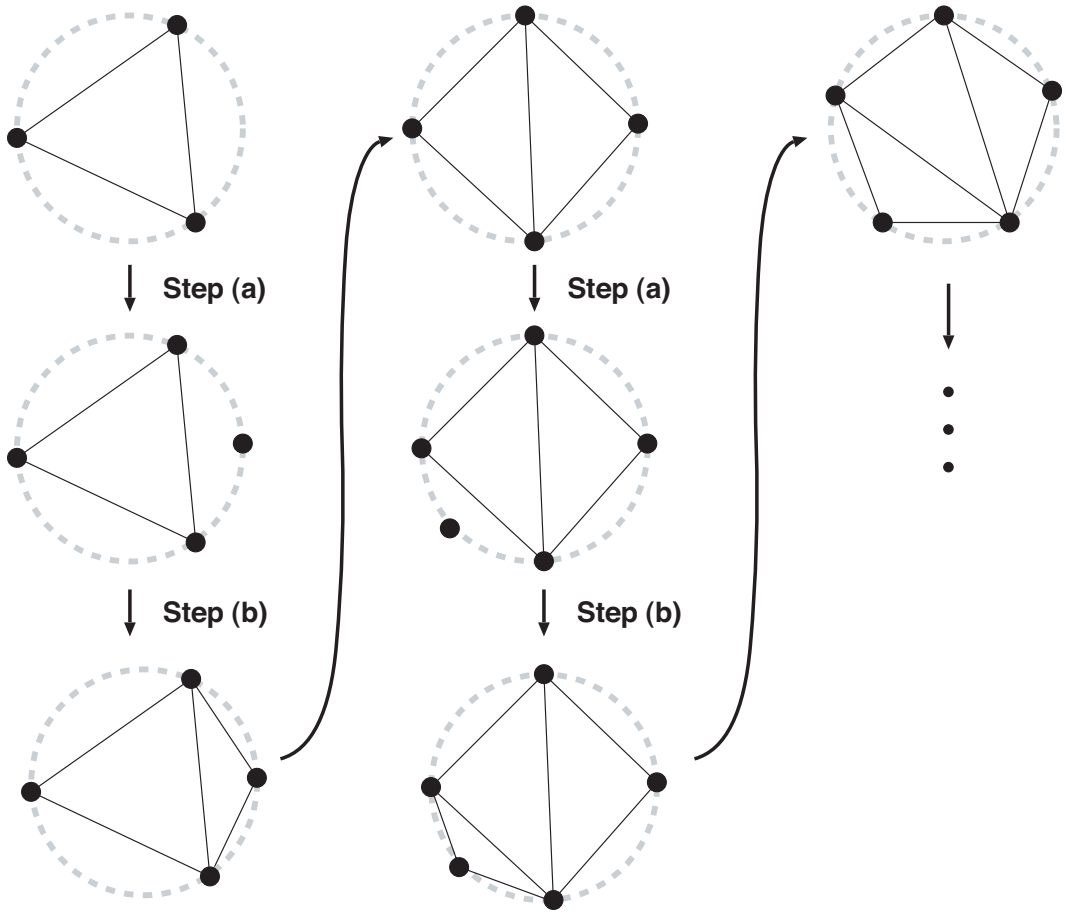


Figure 2.1: Our growing network model, illustrated for $m = 2$. We begin with $m + 1$ completely connected nodes on the circumference of a circle (top left). At each subsequent time step we: a) add a new node in a randomly chosen inter-node interval along the circle circumference, with every interval having equal probability of being chosen, and b) connect the new node to its $m$ nearest neighbors, with nearest here referring to distance along the circle circumference. Steps (a) and (b) are repeated until the desired system size is reached.

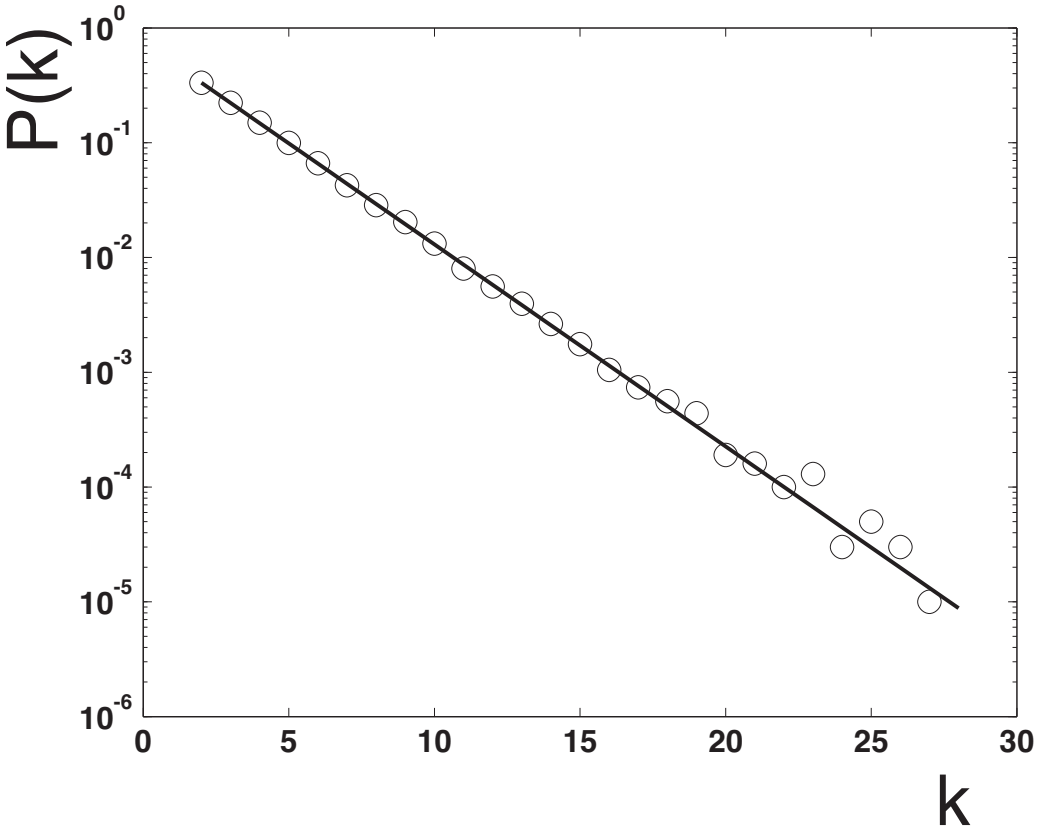


Figure 2.2: The open circles represent the degree distribution $P(k)$ for a network grown according to our model with $N = 10^5$ and the solid line is the analytically calculated ensemble averaged degree distribution [Eq. (2.2)], both with $m = 2$.

second term is the expected number of nodes with degree $k-1$ at time $N$ whose degrees increase to $k$ at time $N+1$. The third term represents the new node with degree $m$.

We let $H(k,N) = G(k,N)/N$ be the fraction of nodes with degree $k$ at time $N$, i.e., the degree distribution. In Appendix A, we show that for large $N$, $H(k,N)$ approaches an asymptotically $N$ invariant form $\bar{H}(k)$, given by

$$\bar{H}(k) = \frac{1}{m+1}\left(\frac{m}{m+1}\right)^{k-m} \tag{2.2}$$

for $k \geq m$ and $\bar{H}(k) = 0$ for $k < m$.

In Fig. 2.2, the data points represent the degree distribution $P(k)$ (i.e., the fraction of nodes with degree $k$) for a single network realization randomly grown by our algorithm (illustrated in Fig.

2.1) for $m = 2$ at $N = 10^5$. The solid line is $\bar{H}(k)$ from Eq. (2.2), also with $m = 2$. We observe good agreement between the analytical calculation for the ensemble average over realizations and the simulation of a single realization, with both showing an exponentially decaying degree distribution. This agreement illustrates that 'self-averaging' applies for large $N$.

In addition, we can calculate the average node degree at time $N$, $\langle k \rangle$, as $N \to \infty$:

$$\underset{\lim N \to \infty}{\langle k \rangle} = \sum_{k=m}^{\infty} k \bar{H}(k) = 2m. \tag{2.3}$$

This can be seen also by observing that each time $N$ increases by 1, $m$ new links are formed, and since each link has two ends, the sum of the degrees of all nodes increases by $2m$ at each time step. Thus our first criterion for a small world network (that $\langle k \rangle$ remains bounded as $N \to \infty$) is met.

## 2.4 Clustering

For the particular case of $m = 2$ we can calculate the average clustering of the network exactly. For this value of $m$, a new node joins the network with $k = 2$ and $q = 1$. Each subsequent addition of a link to that node increments both $k$ and $q$ by one. Thus $q = k - 1$ for all nodes. Since, by definition, $C_i = 2q_i/k_i(k_i - 1)$, the average clustering over all nodes in the $m = 2$ case is given by

$$\langle C \rangle = 2 \left\langle \frac{1}{k} \right\rangle = 2 \sum_{k=2}^{\infty} \frac{1}{k} \bar{H}(k) = \frac{3}{2} \ln 3 - 1 \approx 0.648. \tag{2.4}$$

The open circles in Fig. 2.3 are the node averaged clustering for a single network realization randomly grown by our algorithm (illustrated in Fig. 2.1) versus the network size $N$ for $m = 2$. As $N$ grows, these data are observed to approach the ensemble averaged large $N$ result given by Eq. (2.4) (dashed line). In networks with larger values of $m$ we also observe approach of $\langle C \rangle$ to a constant asymptotic value as $N$ increases (the asymptotic $\langle C \rangle$ grows with $m$; e.g., for $m = 4$, $\langle C \rangle \simeq 0.653$).

The network maintains a high average clustering as $N \to \infty$ and, therefore, the second criterion for a small world network is met. This high clustering is expected due to the local nature of the links made. A new link is inserted in a region that already has high interconnectivity, assuring that the nodes with which the new connections are made have a high probability of

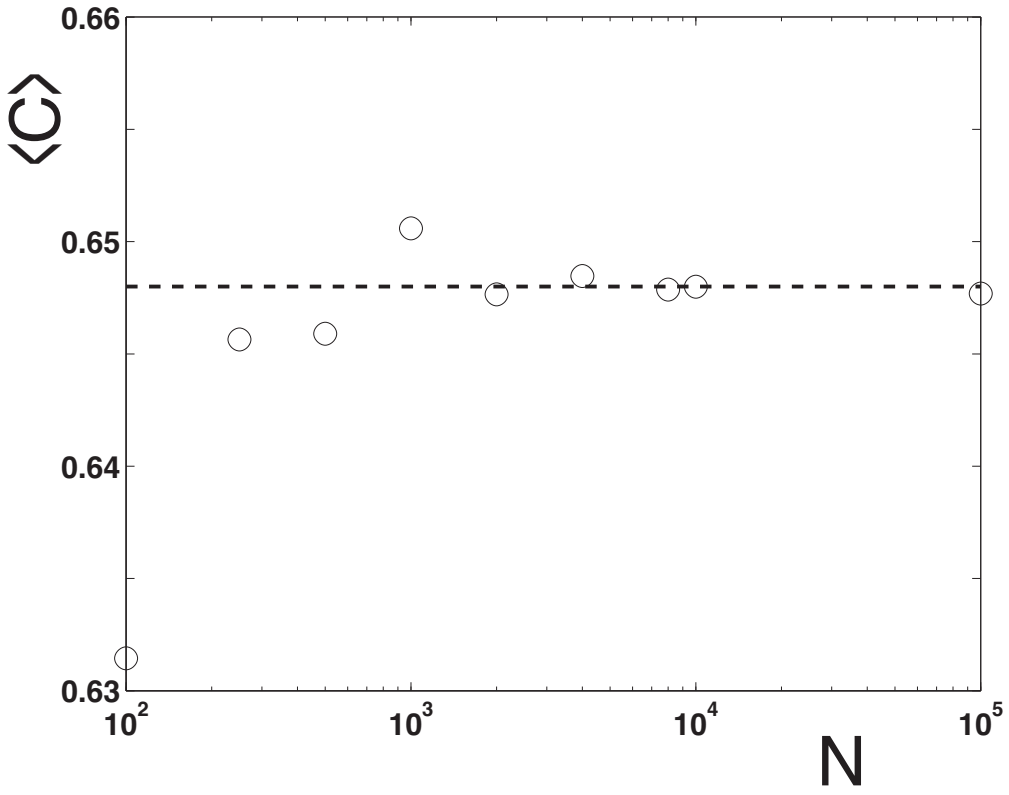


Figure 2.3: Average clustering, $\langle C \rangle$, versus system size, $N$, for a simulated network with $m = 2$. As $N$ grows, the average clustering approaches the value (dashed line) predicted in Eq. (2.4).

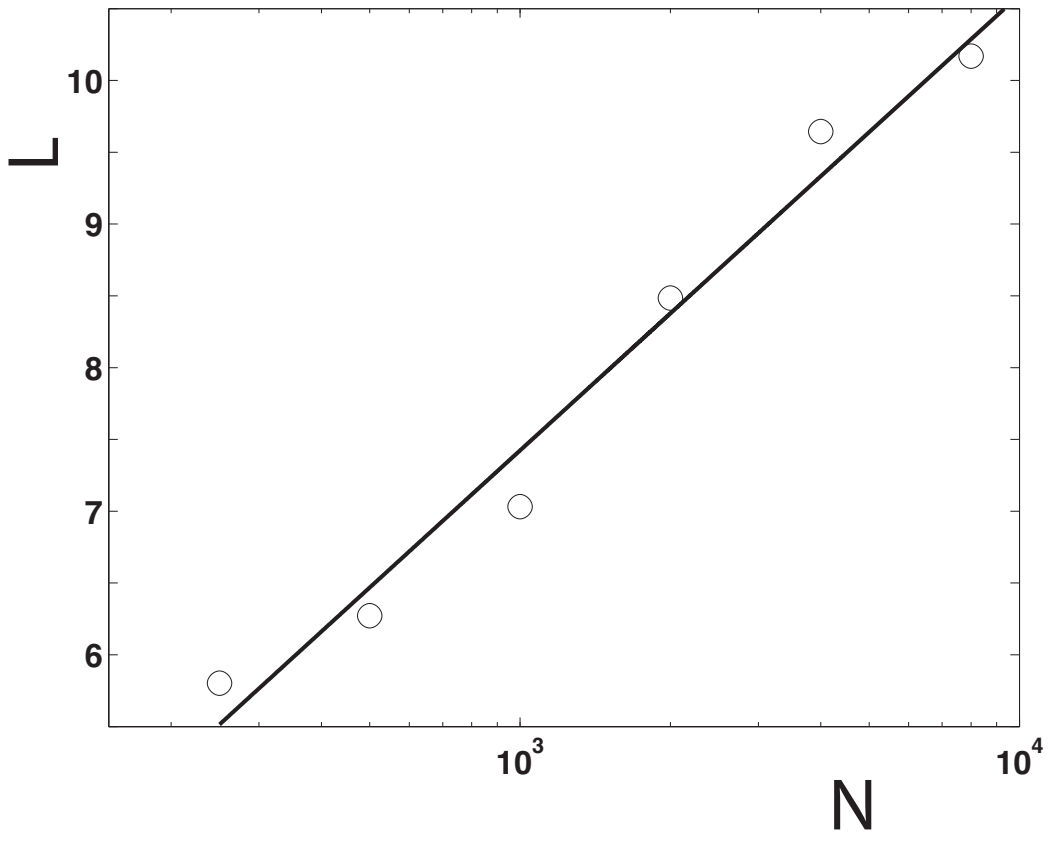


Figure 2.4: Semilogarithmic graph of the characteristic path length, $L$, versus the system size, $N$. The data show the small world slow path length growth characteristic, $L \sim \ln N$. The straight line is a fit to the data.

having connecting links to each other.

## 2.5 Characteristic Path Length

The open circles in Fig. 2.4 show $L$, the shortest path length between pairs of nodes averaged over all node pairs of a single growing network realization, on a linear scale versus $N$ on a logarithmic scale. The data show a linear trend, demonstrating the desired slow growth of geodesic path lengths with system size; i.e. $L \sim \ln N$. Thus the third, and final, small world network criterion is also satisfied.

To see why $L$ grows more slowly than $N$, consider the fact that, although the links made

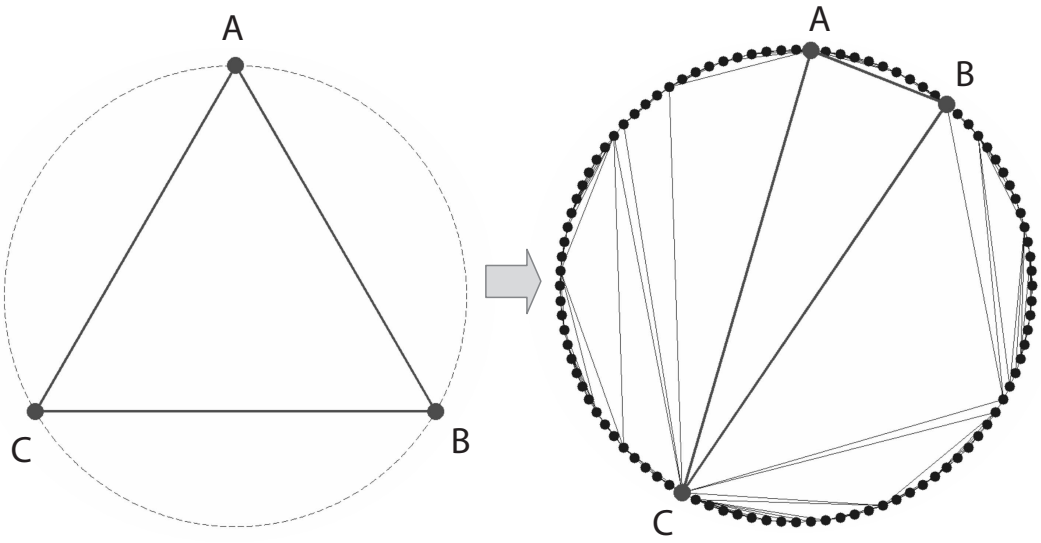


Figure 2.5: An illustration of network growth in our model for $m = 2$. The network starts off (left) with three adjacent nodes, labeled by $A$, $B$, and $C$, connected to each other via links. When the network reaches a network size of 100 nodes (right), the original three labeled nodes are no longer adjacent, but have been "pushed apart" by the new nodes that were inserted between them. The links connecting $A$, $B$, and $C$ serve as short cuts (similar to the shortcuts in the Watts-Strogatz model [4]), resulting in a small characteristic path length for the network.

by incoming nodes are always local, the network itself is growing. The older nodes that had once been nearest neighbors along the circle (and therefore linked), are pushed apart as newer nodes are inserted into the interval between them. Fig. 2.5 illustrates this for the case of $m = 2$. The network begins as three nodes linked to each other. By the time the network reaches $N = 100$, we see that the original nodes are not adjacent, but, rather, have a large number of newer nodes between them. Thus, growth leads to long links between old nodes, and these long links are the shortcuts responsible for a short characteristic path length.

To see why $L \sim \ln N$, imagine a network of size $N \gg 1$ and characteristic path length $L$. Now if we grow the network by adding $N$ new nodes, these nodes will be roughly uniformly distributed along the circle circumference. This means that, on average, a new node would be a

distance of $O(1)$ from one of the first $N$ nodes. Thus as $N \to 2N$ (i.e., $\ln N \to \ln N + O(1)$), we expect $L$ to increase to $L + O(1)$, resulting in $L \sim \ln N$.

## 2.6 Geography

The construction of our network allows for the tracking of the relative positions of the individual nodes. This enables us to ask questions about quantities pertaining to the geographical separation between nodes. By geographical separation we mean that we view our network as equally spaced nodes on the circumference of a circle embedded in a two dimensional Euclidean space and this allows us to define a physical separation between the nodes. For convenience, we consider the distance between nodes as the distance measured along the circumference of the circle in units of the inter-node spacing or, in other words, one plus the number of nodes found on the circle's circumference between the two nodes of interest[1]. (This is different from the graph theoretic notion of the distance between two nodes defined as the minimum number of links connecting two nodes.)

Here we present some analytical results related to geographical distances in our network for the case $m = 2$. Let $\hat{S}^+(k, l, N)$ be the number of nodes at time $N$ with degree $k$ and with the longest link to another node in the counter-clockwise (CCW) direction having the length $l$. (A node $i$ has its longest CCW link connected to the node $j$ which was the nearest neighbor in the CCW direction when node $i$ was introduced into the network as a new node.) At each time step a new node is added to the network and it connects to its two nearest neighbors which are, by definition, one inter-node spacing away from the new node. Thus the new node will have $k = 2$ and $l = 1$. As more nodes are added, the values of $k$ and $l$ for an existing node will change according to the following scenarios. Figure 2.6 is an illustration of a section of our network, depicting the possible scenarios for the node labeled $B$. Let node $B$ have degree $k$ and CCW longest link $l$. It is clear that if a new node is introduced in region IV, neither $k$ nor $l$ is affected. If a new node is placed anywhere in region III, defined as the region between the CCW nearest neighbor node $C$ and the farthest CCW linked node $D$, then $l$ increases by one while $k$ remains the same. A new

[1] We do not use the Euclidean distance between nodes, which would be the length of the chord connecting two nodes on the circle's circumference.

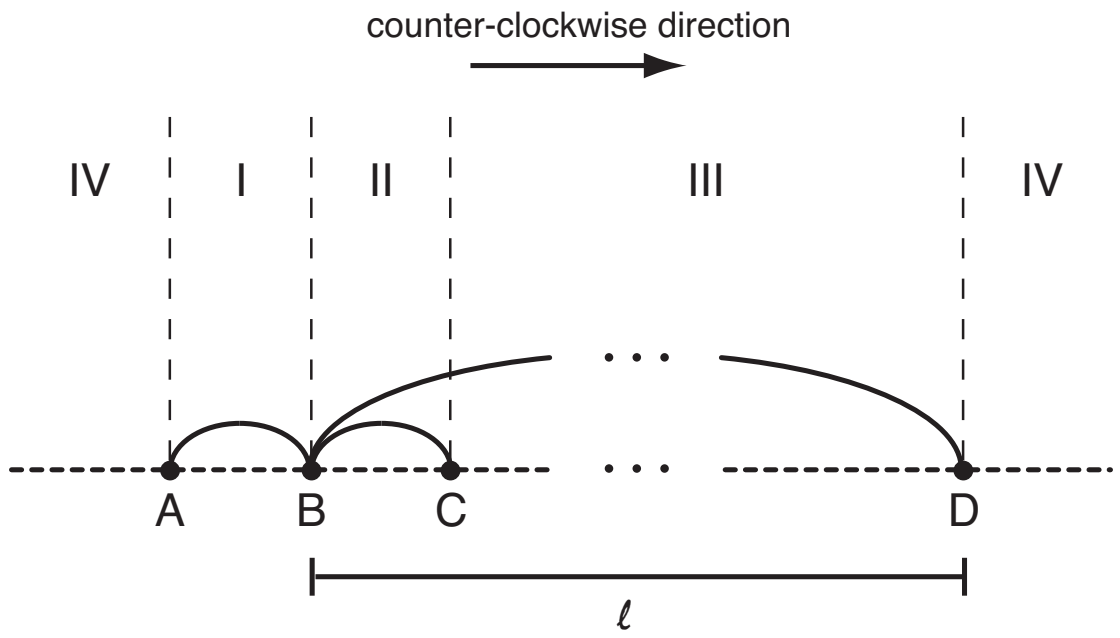


Figure 2.6: An illustration of the different regions associated with node $B$. Let node $B$ have degree $k$ and longest CCW link $l$. A new node inserted in region IV will not affect either $k$ or $l$. A new node inserted in region III (between the CCW nearest neighbor $C$ and the CCW farthest connected node $D$) increases $l$ by one but leaves $k$ unchanged. A new node inserted in region II (between nodes $B$ and $C$) increases both $k$ and $l$ by one. Since region I (between nodes $A$ and $B$) is in the clockwise direction, a new node inserted here increases $k$ by one but leaves $l$ unchanged.

node inserted into region II increments both $k$ and $l$ by one. Finally, since we are only interested in the longest CCW (as opposed to clockwise) link of node $B$, if a new node is inserted in region I, $k$ is incremented by one while $l$ remains unchanged.

We now take the average over all realizations of the possible random placements of the new node and write down the following evolution equation for the average of $\hat{S}^+$, which we denote $S^+$:

$$\begin{aligned} S^+(k,l,N+1) &= \left(1-\frac{l+1}{N}\right) S^+(k,l,N) + \frac{1}{N}S^+(k-1,l,N) \\ &\quad + \frac{1}{N}S^+(k-1,l-1,N) + \frac{l-2}{N}S^+(k,l-1,N) + \delta_{k2}\delta_{l1}. \end{aligned} \tag{2.5}$$

The first term on the right hand side is associated with region IV and represents the expected number of nodes with degree $k$ and longest CCW link $l$ at time $N$ which will retain those values

at time $N+1$. The second term is the expected number of nodes with degree $k-1$ and longest CCW link $l$ that receive a new node in region I at time $N$ and therefore increase their degree to $k$ while leaving $l$ unchanged at time $N+1$. The third term is the expected number of nodes with degree $k-1$ and longest CCW link $l-1$ at time $N$ which receive a node in region II. The fourth term is the expected number of nodes with degree $k$ and longest CCW link $l-1$ at time $N$ which receive a node in region III. Note that region III is of length $l-2$ in this case. The last term in the equation accounts for the new node ($k=2$, $l=1$) that is introduced at time $N$.

Defining $s^+(k,l,N) = S^+(k,l,N)/N$ to be the fraction of nodes with degree $k$ and longest CCW link $l$ at time $N$, we can write down the evolution equation for $s^+$ as,

$$\begin{aligned}(N+1)s^+(k,l,N+1) &= (N-(l+1))s^+(k,l,N) + s^+(k-1,l,N) \\ &\quad + s^+(k-1,l-1,N) + (l-2)s^+(k,l-1,N) + \delta_{k2}\delta_{l1}.\end{aligned} \tag{2.6}$$

We define $W(k,N) = \sum_l s^+(k,l,N)$ to be the fraction of nodes with degree $k$ at time $N$. Thus $W(k,N)$ should coincide with $H(k,N)$ for $m=2$ (see Sec. 2.3). This is indeed the case, as summing both sides of Eq. (2.6) over $l$ recovers an evolution equation for $W(k,N)$ identical to the evolution equation for $H(k,N)$ with $m=2$ (i.e., Eq. (A.1) with $m=2$). Hence, the stationary ($N$ independent) solution $\bar{W}(k)$ of $W(k,N)$ satisfies the recursion relation

$$\bar{W}(k) = \frac{2}{3}\bar{W}(k-1) + \frac{\delta_{k2}}{3}, \tag{2.7}$$

with the solution

$$\bar{W}(k) = \frac{1}{3}\left(\frac{2}{3}\right)^{k-2}. \tag{2.8}$$

We now introduce the quantity $\langle l|k\rangle_N^+ = \sum_l l s^+(k,l,N)/W(k,N)$, which is the average longest CCW link for a node with degree $k$ at time $N$. Returning to Eq. (2.6), we multiply both sides by $l$ and then sum over $l$, and obtain the following evolution equation:

$$\begin{aligned}(N+1)\langle l|k\rangle_{N+1}^+ W(k,N+1) &= (N-1)\langle l|k\rangle_N^+ W(k,N) \\ &\quad + 2\langle l|k-1\rangle_N^+ W(k-1,N) \\ &\quad + W(k-1,N) - W(k,N) + \delta_{k2}.\end{aligned} \tag{2.9}$$

If we assume the stationary solutions $\langle l|k\rangle^+$ and $\bar{W}(k)$ for $\langle l|k\rangle_N^+$ and $W(k,N)$, respectively, we are left with the recursion relation

$$\bar{W}(k)\left(2\langle l|k\rangle^+ + 1\right) = \bar{W}(k-1)\left(2\langle l|k-1\rangle^+ + 1\right) + \delta_{k2}, \tag{2.10}$$

which, using the result in Eq. (2.7), can be solved to yield

$$\langle l|k\rangle^+ = \left(\frac{3}{2}\right)^{k-1} - \frac{1}{2} \tag{2.11}$$

for $k \geq 2$ and $\langle l|k\rangle^+ = 0$ for $k < 2$. Thus we see that the length of the longest CCW link of a node (as well as its longest clockwise link, by symmetry) increases exponentially with its degree. We show in Appendix B that the average longest link of a node also behaves in a similar manner.

Turning our attention to the distribution of the longest CCW link lengths, we define the quantity $T(l,N) = \sum_k s^+(k,l,N)$, the fraction of nodes with longest CCW link $l$ at time $N$. Summing both sides of Eq. (2.6) over $k$ we obtain the evolution equation

$$(n+1)T(l,N+1) = (N-l)T(l,N) + (l-1)T(l-1,N) + \delta_{l1}. \tag{2.12}$$

We define the stationary solution $\bar{T}(l)$ of $T(l,N)$, and from Eq. (2.12) obtain the recursion relation

$$\bar{T}(l)\,(l+1) = \bar{T}(l-1)\,(l-1) + \delta_{l1} \tag{2.13}$$

which we solve to yield

$$\bar{T}(l) = \frac{1}{(l+1)\,l} \tag{2.14}$$

for $l \geq 1$ and $\bar{T}(l) = 0$ for $l < 1$.

We now define the quantity $\langle k|l\rangle_N^+ = \sum_k k\, s^+(k,l,N)/\,T(l,N)$, the average degree for a node with longest CCW link $l$. We multiply (2.6) by $k$ and sum over $k$, obtaining the evolution equation

$$\begin{aligned}(N+1)\,\langle k|l\rangle_{N+1}^+\,T(l,N+1) \;=\;& (N-l)\,\langle k|l\rangle_N^+\,T(l,N) \\ &+ (l-1)\,\langle k|l-1\rangle_N^+\,T(l-1,N) \\ &+ T(l,N) + T(l-1,N) + 2\delta_{l1}.\end{aligned} \tag{2.15}$$

Assuming the stationary solutions $\langle k|l\rangle^+$ and $\bar{T}(l)$ for $\langle k|l\rangle_N^+$ and $T(l,N)$, respectively, we obtain the recursion relation

$$\bar{T}(l)\left((l+1)\langle k|l\rangle^+ - 1\right) = \bar{T}(l-1)\left((l-1)\langle k|l-1\rangle^+ + 1\right) + 2\delta_{l1}, \tag{2.16}$$

which, using the result in (2.13), is solved to yield

$$\langle k|l\rangle^+ = \frac{5}{2} + \sum_{i=2}^{l} \frac{2i}{i^2-1} \tag{2.17}$$

for $l \geq 1$ and $\langle k|l\rangle^+ = 0$ for $l < 1$. In particular, this implies that for large $l$, $\langle k|l\rangle^+ \sim \ln l$.

## 2.7 Conclusion

We presented a small world network model that has only geographically local interactions. This model provides a physically realistic mechanism by which growing physical systems that have geographical constraints, and therefore limited global information available to each individual node, can form networks with small world characteristics. Additionally, our results suggest that small world networks in geographically constrained physical systems may be a natural consequence of system growth and local interactions.

Chapter 3

# Formation of Multifractal Population Patterns from Reproductive Growth and Local Resettlement

## 3.1 Introduction

Growing populations exist in many different areas of interest. Although the most obvious examples are biological in nature (e.g., bacterial cultures, the human population), one can also consider the growth of technological development or the growth of urban infrastructure as more abstract examples. While the specific details of these systems can be quite different, the examples stated here all share the characteristic that their populations grow on some background space. A reasonable question to ask then is what types of spatial distributions result in these growing systems. Figure 3.1 is a version of the popular Earth at Night image (EaN) [3]. This is a composite image taken by orbiting satellites that shows the Earth's nighttime man-made lights as seen from space. The light intensities are brightest in areas that are known to be highly populated and developed, and thus the image can be crudely thought of as representing the spatial distribution of some combined measure of technological development and human population density. We see a very heterogeneous distribution, with areas of very high intensities as well as areas of almost no light at all.

A useful characterization of the heterogeneity of a distribution is the distribution's *dimension spectrum* $D_q$. A grid of hypercubes, or boxes, of a fixed size with edge length $\epsilon$ is used to cover the distribution, where the dimension of the hypercubes corresponds to the dimension of the underlying space (e.g., squares for two dimensions and line segments for one dimension). $D_q$ is then calculated via [18, 19]

$$D_q = \frac{1}{1-q} \lim_{\epsilon \to 0} \frac{\ln I(q,\epsilon)}{\ln(1/\epsilon)}, \tag{3.1}$$

where

$$I(q,\epsilon) = \sum_{i=1}^{N(\epsilon)} \mu_i^q. \tag{3.2}$$

Here the sum is taken over $N(\epsilon)$ non-empty grid boxes and $\mu_i$ is the fraction of the particular quantity of interest (assumed to be nonnegative, e.g., light intensity in Fig. 3.1) that is contained

Figure 3.1: A version of the Earth at Night image [3]. This is a composite image created with data from the Defense Meteorological Satellite Program of the Earth's nighttime man-made lights as seen from space. For the analysis, an 8-bit grayscale 2400x800 pixel image was used, where light from the reflective landmasses found in the original image was removed with an appropriate uniform background subtraction. The image pixels have intensity values ranging from 0 to 218.

in box $i$.

The parameter $q$ ($q \geq 0$ in this chapter) can be varied continuously to vary the influence of high and low $\mu_i$ boxes, with larger values of $q$ emphasizing the boxes with larger $\mu_i$'s. In particular, when $q = 1$ we have the *information dimension* $D_1$ which, applying L'Hospital's Rule to (3.1) and (3.2), is given by

$$D_1 = \lim_{\epsilon \to 0} \frac{\sum_{i=1}^{N(\epsilon)} \mu_i \ln \mu_i}{\ln \epsilon}. \tag{3.3}$$

A distribution is *fractal* if it possesses a fractional (non-integer) dimension (i.e., if $D_q$ is not an integer for some $q$). If $D_q$ depends on $q$, then the distribution is said to be *multifractal*.

Generally, when measuring the dimension of a distribution of numerical and/or experimental data, there is a maximum resolution observable. In the EaN image, for example, the image's individual pixels are the smallest boxes that can be used to cover the distribution in the calculation of $D_q$. In such cases the $\epsilon \to 0$ limit cannot be considered; instead one looks for a *scaling range* of $\epsilon$ where the quantity $(1-q)^{-1} \ln I(q, \epsilon)$ varies approximately linearly with $\ln(1/\epsilon)$, from which a dimension can be extracted as the slope of a straight line fitted to data in the $\epsilon$ scaling range. We illustrate this process by calculating the information dimension $D_1$ of the EaN image. Figure 3.2

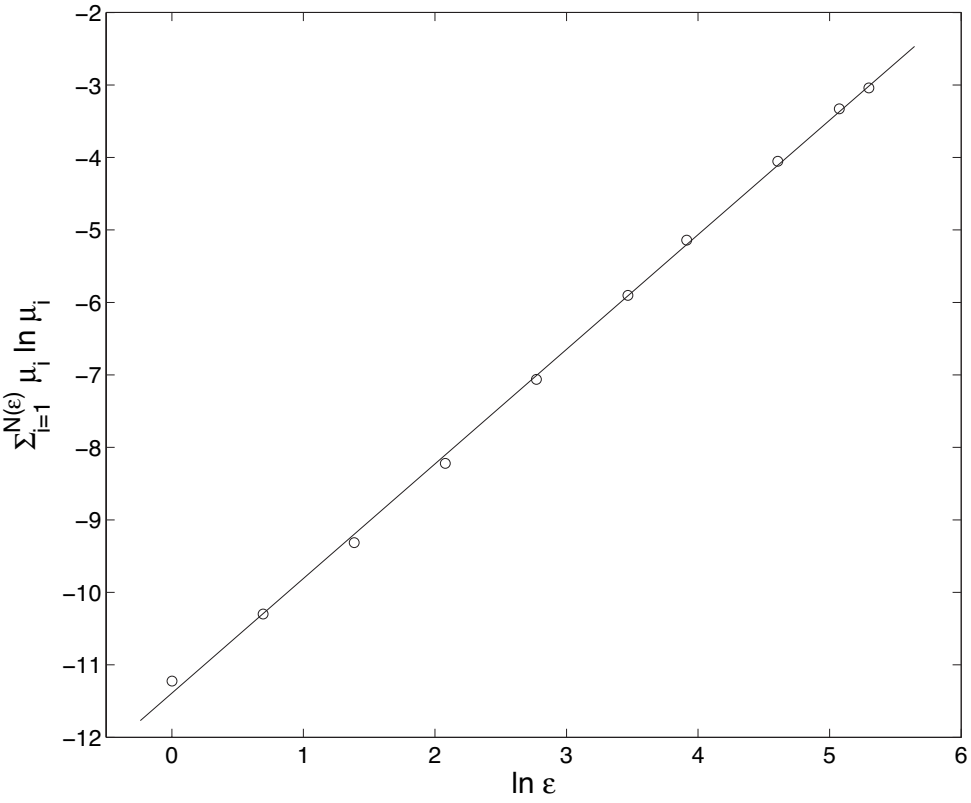


Figure 3.2: Plot for calculating the information dimension $D_1$ [see Eq. (3.3)] for the Earth at Night image. The quantity $\sum_{i=1}^{N(\epsilon)} \mu_i \ln \mu_i$ is plotted versus $\ln \epsilon$ (open circles), where $\epsilon$ is in units of pixels, resulting in a linear *scaling region*, the slope of which is a finite-scale approximation of $D_1$. The solid line is a linear fit to the data with a slope of $1.60 \pm 0.02$.

shows a plot of $\sum_{i=1}^{N(\epsilon)} \mu_i \ln \mu_i$ versus $\ln \epsilon$ (with $\epsilon$ in units of pixels) for the image, where each $\mu_i$ is the fraction of the total intensity in box $i$. The good evidence of linearity in this plot over a scaling range in $\epsilon$ of order $e^6$ indicates that a fractal description makes sense. The solid line in Fig. 3.2 is a linear fit to the data with a slope of $1.60 \pm 0.02$, which we take as the value of $D_1$. Thus the light distribution in the EaN is fractal. In Sec. 3.3 we argue that it is also multifractal.

Yook et al. [20] have recently reported that the world's human population, as well as the population of internet routers and autonomous systems, are fractally distributed. Additionally, several studies [21–23] have demonstrated various fractal properties of urban settlements. We

were thus led to consider the possibility of common, underlying mechanisms shared by growing populations that lead to fractal distributions.

One characteristic found in many growing populations is the existence of some generalized reproduction process, where existing members of the population generate more members. In addition to biological reproduction, in some non-biological systems concentrated populations can encourage the creation of more members. For example, areas that are technologically highly developed are likely to stimulate more development than areas that are not as technologically developed.

Once the new members are generated, it is reasonable to expect some form of local resettlement for a wide range of growing systems. The resettlement is necessary in populations where individual members are unable to occupy the same physical space as other members.

In this chapter we demonstrate that these two ingredients, reproduction and local resettlement, can lead to multifractal spatial distributions for growing populations. We present several point placement models in one and two dimensions which implement these ideas and show, both numerically and analytically, that they lead to multifractal distributions.

One might suspect that fractal population densities arise due to strong inhomogeneities of the underlying space (e.g., inhomogeneities in the distribution of land fertility, natural resources, etc.). Our models initially (Sec. 3.2) employ spaces that have *no inhomogeneity*, thus demonstrating that fractal population distributions can occur without inhomogeneity. We also consider generalizations to our models that include geographical inhomogeneities (Sec. 3.4) and demonstrate that, for the particular forms of the inhomogeneities that we employ, the multifractality of the distribution is unaffected.

In Sec. 3.2 we present our point placement models along with numerical results. (We emphasize that the set of models we introduce in Sec. 3.2 is not exhaustive, and many other similar models could be conceived.) Section 3.3 considers situations where a multifractal distribution is sensed by an instrument that saturates at a maximum measurable value (clipping). We show that such clipping occurs in the EaN data, affecting the determined $D_q$. In Section 3.4, we investigate

the effect of adding geographical inhomogeneity to the underlying space of our models. In Sec. 3.5 we present analytical results for some of our models. We summarize our findings and conclude in Sec. 3.6. Appendix C provides details and background on the analytical calculation of the spectra of fractal dimensions for some of our models. Appendix D contains the proof of a theorem used in Appendix C (this theorem should be of very general utility, not restricted to our specific models, in analysis of the spectra of fractal dimensions).

## 3.2 Models and Results

### 3.2.1 1-D Random Interval (Model 1)

We begin with a very simple model that places points on the unit interval ($0 \leq x \leq 1$). Our initial state is a single point at $x = 0.5$. At each subsequent discrete time step we place a new point according to the following prescription: (a) a target point, or "parent", is chosen from all of the preexisting points with equal probability; (b) the new point, or "child", is placed in the middle of one of the two empty intervals adjacent to the target point with equal probability. Steps (a) and (b) are repeated until the desired system size is reached. Several representative steps of this construction are illustrated in Fig. 3.3.

In Step (a), by randomly choosing the target point from all available preexisting points with equal probability, we allow every point to generate, attract, or reproduce new members at the same rate. This also implies that areas with a high density of points are more likely to attract new members than areas that are sparsely populated, creating a "rich get richer" phenomenon. We will see that all the models presented in this chapter have this general characteristic, with most of the models actually possessing the identical Step (a). Step (b), on the other hand, is the local resettlement scheme that is particular to this model. It was chosen for its simplicity, and we will show in Sec. 3.5.1 that analytical results can be obtained for it.

Figure 3.4 shows semilog plots of point location histograms with different bin sizes $\Delta x$ for one realization of our point placement scheme. We observe a roughening of the plot as smaller histogram bins are used, thus giving a sense of the heterogeneity of the distribution. To quantify

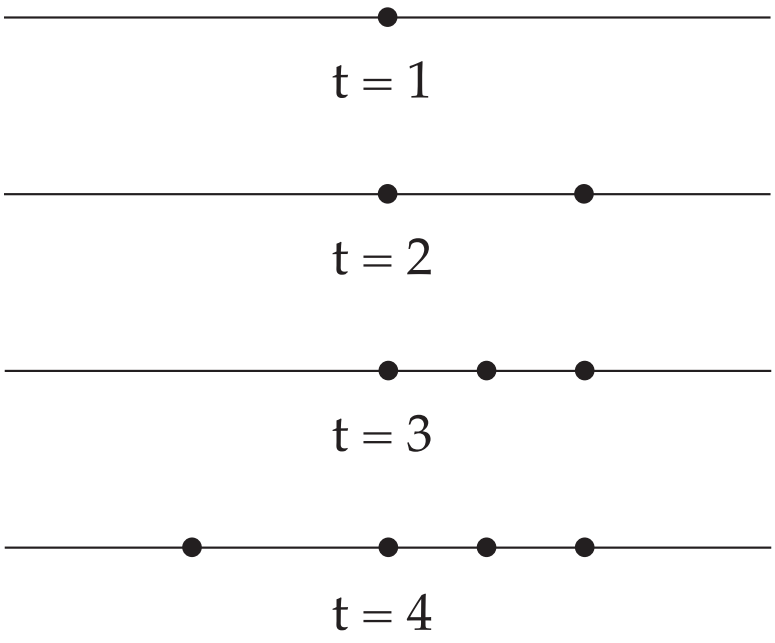


Figure 3.3: Three representative steps of Models 1 and 3.

this heterogeneity, we calculate the dimension spectrum $D_q$ of the distribution. We cover the distribution of points with intervals of a fixed length $\epsilon$ and assign to each interval $i$ a measure $\mu_i$ equal to the fraction of the total points contained in $i$. Then, referring to Eq. (3.1), we plot the quantity $(1-q)^{-1} \ln I(q,\epsilon)$ versus $\ln(1/\epsilon)$ for a range of $\epsilon$, so that the slope of the graph in the scaling range gives us $D_q$.

For $q \geq 0.5$ the plots exhibit a linear regime that allows for a relatively unambiguous determination of the dimension. However, for the smaller $q$ there tend to be shorter scaling regions, which makes the extraction of a reliable $D_q$ more ambiguous. This is due to the fact that the distribution is made up of a finite collection of points. As the grid covering the distribution is made finer, grid boxes in the more sparsely populated areas begin to contain either 0 or 1 points. Furthermore, if $\epsilon$ is made small enough, eventually all the points in the distribution are covered by different boxes, and the dimension for the distribution is simply the dimension of a collection of individual points, i.e., zero. Thus, in order to get a meaningful dimension out of a finite distribution of points, a certain degree of coarseness of the covering grid is necessary. For small values of $q$, $D_q$ relies heavily on boxes containing small amounts of the measure. Therefore,

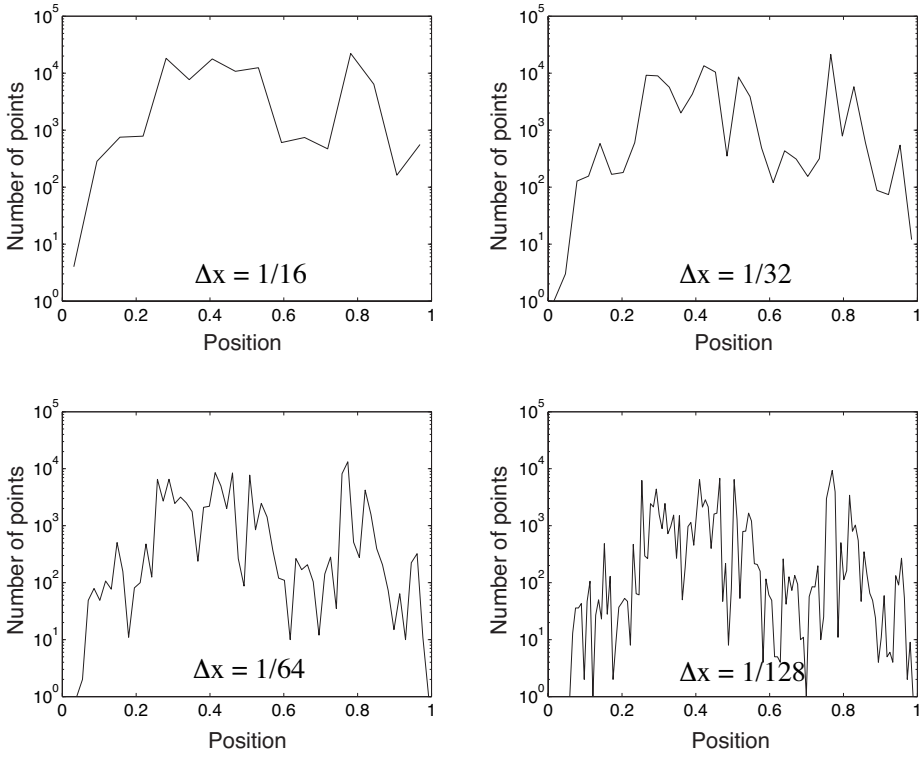


Figure 3.4: Semilog plots of histograms of point locations on the unit interval for a $10^5$ point distribution generated by Model 1, for various bin sizes $\Delta x$. a) $\Delta x = 1/16$, b) $\Delta x = 1/32$, c) $\Delta x = 1/64$, d) $\Delta x = 1/128$.

as $\epsilon$ is decreased, a scale is quickly reached where points in sparsely populated areas are covered by separate boxes, thereby destroying the overall scaling behavior. For larger $q$, the dense regions have a large enough influence on $D_q$ that the effects of the finiteness of the distribution are not observed until very small scales are reached, allowing for a large scaling range.

Figure 3.5 is a plot of the numerically determined average $D_q$ (open circles) for 20 realizations of $10^6$ point distributions generated by Model 1. The solid line represents the analytical result for $D_q$ derived in Sec. 3.5.1. Both the numerical and analytical results demonstrate that Model 1 generates multifractal distributions.

Notice that our analytical value for the box-counting dimension $D_0$ is one. Unlike $D_q$ for $q$ positive, $D_0$ depends only on the limiting set of points and not their distribution. Our result that $D_0 = 1$ indicates that the set filled by the points in Model 1 as time $t$ goes to infinity is not fractal. In fact, the entire interval is filled with probability one, which we see as follows. At time $t$ there are $t$ points, which divide the unit interval into $t+1$ subintervals (see Fig. 3.3). Each of the subintervals has probability $1/t$ of containing the next point, except for the two end subintervals, which each have probability $1/(2t)$. Thus the probability that a given subinterval is not bisected at time $t$ is at most $1 - 1/(2t)$. Since the infinite product of these probabilities as $t \to \infty$ is zero, each subinterval is eventually bisected with probability one, and therefore no empty intervals remain as $t \to \infty$. One can similarly argue for the other models in this chapter that the limiting sets are not fractal; the heterogeneous point distributions they yield are reflected by the fractional values of $D_q$ for $q > 0$.

### 3.2.2 2-D Random Square (Model 2)

Model 2 is a two dimensional analogue of Model 1, where points are placed on the unit square, $(0 \le x \le 1,\ 0 \le y \le 1)$, instead of the unit interval. We begin with a single point at $(0.5, 0.5)$. Imagining vertical and horizontal lines drawn through this point, we see that the point is at the vertex of four neighboring squares. At the next step, we choose one of the squares at random, with equal probability, and place a point in its middle. We then divide this chosen square into four

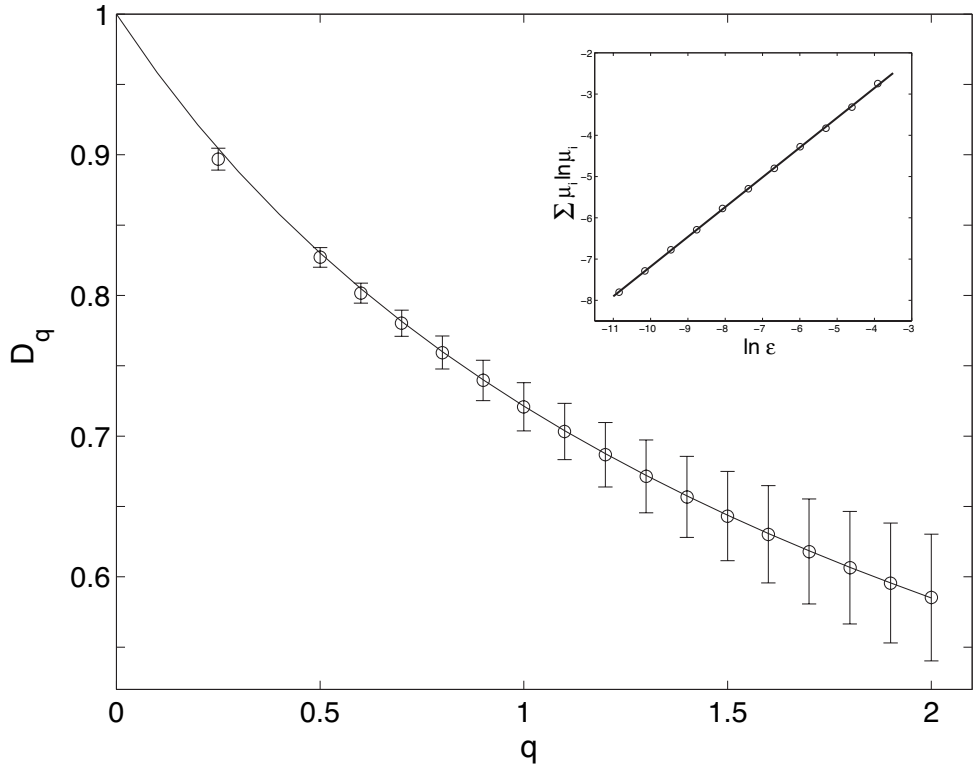


Figure 3.5: Plot of $D_q$ for Model 1. The open circles are the numerical values calculated from the slope of a line fit to $(1-q)^{-1} \ln I(q,\epsilon)$ versus $\ln(1/\epsilon)$ [see Eq. (3.1)], averaged over 20 realizations of $10^6$ point distributions (except for $q = 1$, in which case the line is fit to $\sum_{i=1}^{N(\epsilon)} \mu_i \ln \mu_i$ versus $\ln \epsilon$ [see Eq. (3.3)]), with the error bars indicating the sample standard deviation for each value. The solid line is the analytical result of Eq. (3.14). As an example the inset shows the determination of $D_1$ for one point distribution.

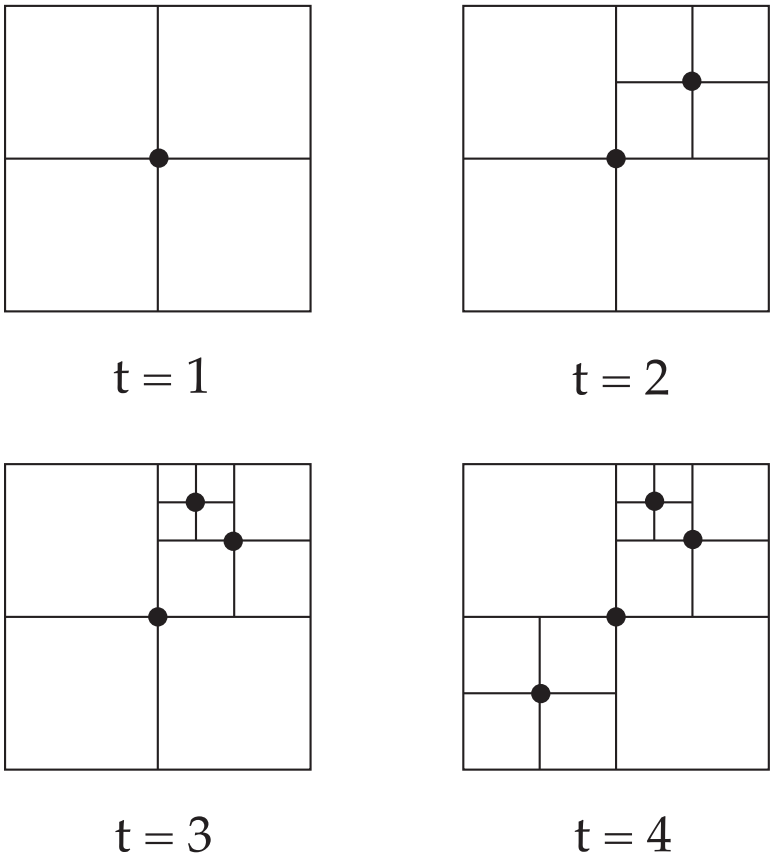


Figure 3.6: Three representative steps of Models 2 and 4.

smaller squares by horizontal and vertical lines through the newly added point. Then we continue this process: At each subsequent discrete time step we pick a target point (parent) according to Step (a) of Model 1 (that is, all points are equally likely to be chosen) and add a new point (child) to the center of one of the target point's four adjacent squares, chosen with equal probability. Several representative steps of this construction are illustrated in Fig. 3.6.

Figure 3.7 shows a distribution generated by this model's point placement scheme, where we have successively magnified dense regions of the distribution to illustrate the large differences in densities that result at various length scales. Figure 3.8 is a plot of the average $D_q$ of 20 independent distributions generated by this model, where the solid line is the analytical result (3.28) derived in Sec. 3.5.2. Again, both the numerical and analytical results confirm that Model 2 generates multifractal distributions.

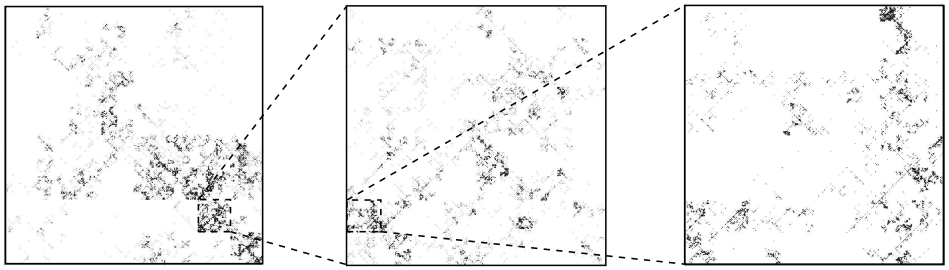

Figure 3.7: A plot of a $4 \times 10^5$ point distribution on the unit square generated by Model 2. Dense regions are magnified to illustrate the heterogeneity of the point densities in the distribution at different scales.

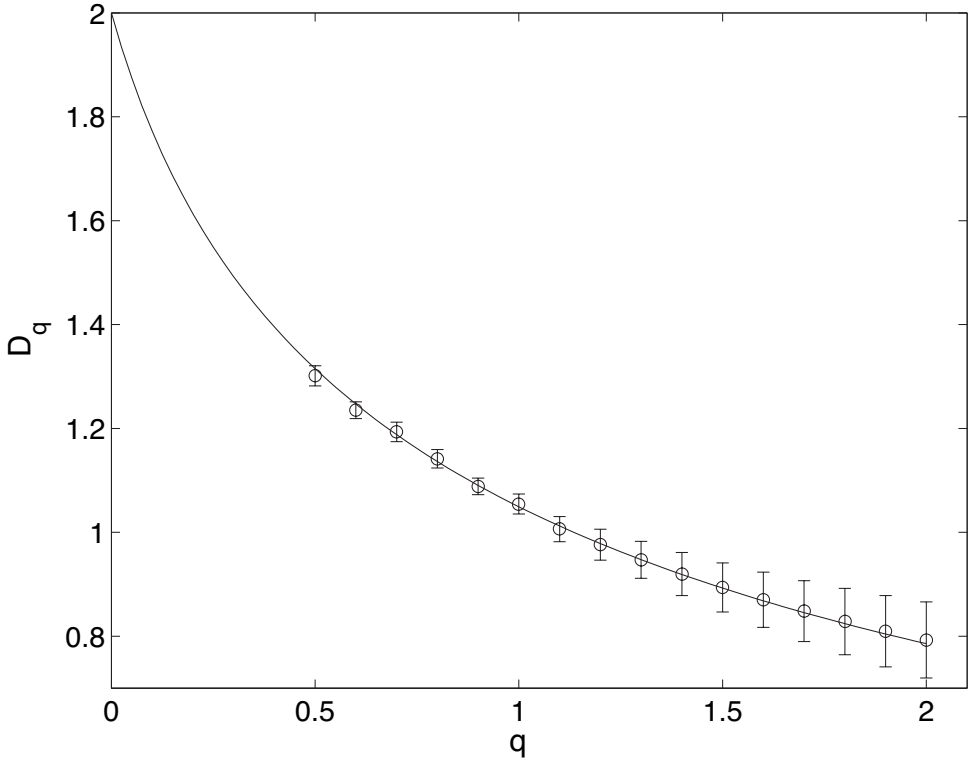


Figure 3.8: Plot of $D_q$ for Model 2. The open circles are the average numerical values of $D_q$ (see caption in Fig. 3.5) obtained from 20 realizations of $10^6$ point distributions generated by this model and the error bars are the sample standard deviations for each value. The solid line is the analytical result of Eq. (3.28).

### 3.2.3 1-D Larger Interval (Model 3)

Model 3 is a one dimensional point placement model similar to Model 1, but with a different local resettlement rule. That is, the initial state and Step (a) are identical to those of Model 1, while Step (b) is modified. Instead of a new point being randomly placed in one of the adjacent intervals of the chosen target point with equal probability, the new point is placed in the larger of the two intervals adjacent to the target point, and, if the two intervals are equal in size, one is chosen at random.

In this way, the new points seek out more sparsely populated local intervals to settle within. This strategy is reasonable in a context such as human resettlement, where the settlers may try to combine the convenience of being close to others with the possible advantages associated with more space. We expect this resettlement scheme to yield distributions that are more homogeneous than those created by Model 1, and hence yield greater $D_q$.

We plot the point location histograms for one distribution generated by this model in Fig. 3.9, once again varying the bin sizes used. Although the roughening of the plot as the bin size is decreased is less pronounced than what was observed in Fig. 3.4 for Model 1, there still appear to be significant fluctuations in the densities. We numerically obtain $D_q$ and find the dimension spectrum to be multifractal, having the same qualitative shape as the $D_q$ for Model 1. However, due to the more uniform nature of the distributions created by this model, the specific values of $D_q$ are larger at each $q$ (e.g., $D_1 = 0.84 \pm 0.01$ [averaged over 20 distributions of $10^6$ points] for this model versus our analytical result $D_1 \simeq 0.72$ for Model 1 [Eq. 3.15]).

### 3.2.4 2-D Sparse Square (Model 4)

Model 4 is a two dimensional analogue of Model 3. The initial state and Step (a) of its construction are the same as in Model 2. Our aim is to extend Step (b) of Model 3, the choosing of the larger interval for settlement, to two dimensions and for that purpose we make the following observation. In the square-based two dimensional point placement scheme of Model 2, there are two different types of squares which we refer to as type I and type II squares. A type I square only has one

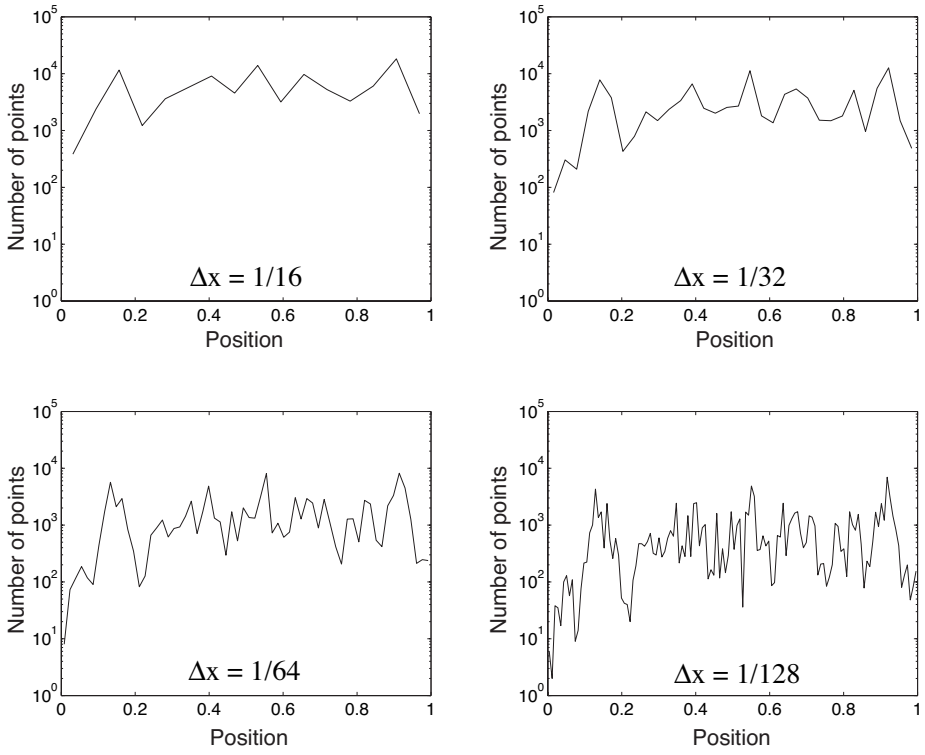


Figure 3.9: Semilog plots of histograms of point locations on the unit interval for a $10^5$ point distribution generated by Model 3, for various bin sizes $\Delta x$. a) $\Delta x = 1/16$, b) $\Delta x = 1/32$, c) $\Delta x = 1/64$, d) $\Delta x = 1/128$.

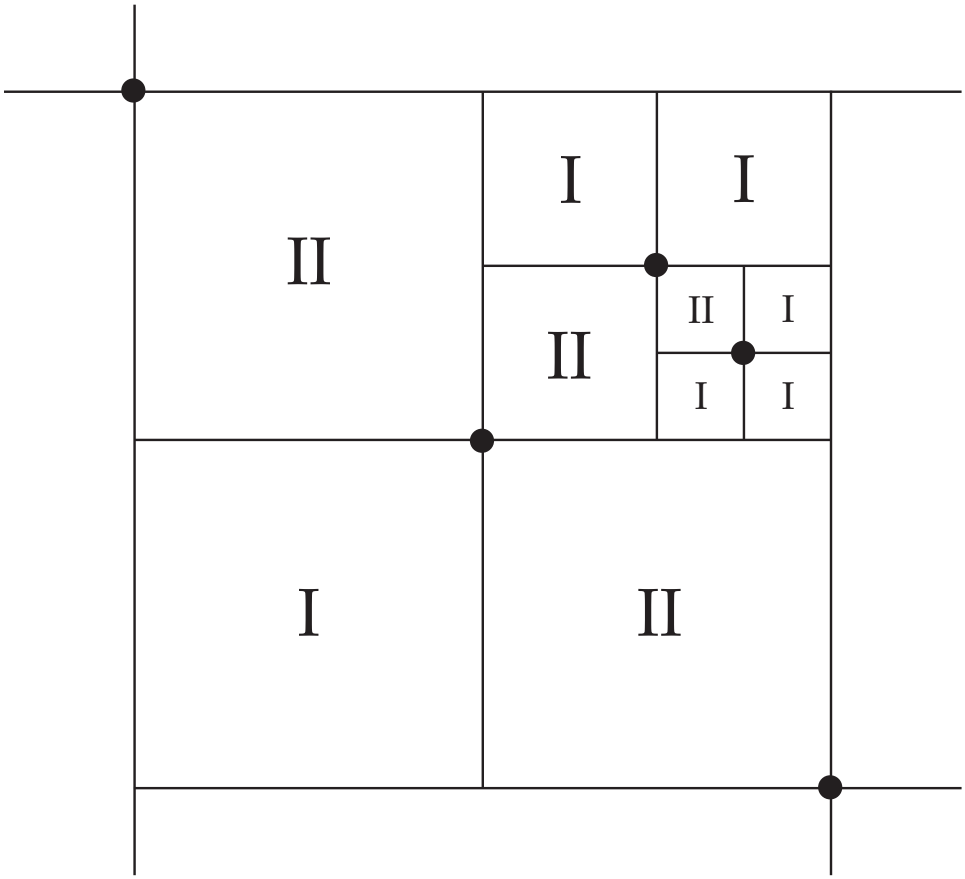


Figure 3.10: Illustration of type I (labeled I) and type II (labeled II) squares, where black dots indicate populated vertices.

of its vertices occupied by a point, while a type II square has two opposite vertices occupied by points. (See Fig. 3.10 for an illustration of type I and type II squares.) Thus a type I square can be regarded as more sparsely populated than a type II square of equal size. With this in mind, Step (b) for Model 4 is to place the new node in the middle of the largest square adjacent to the target point. In case of a tie, choose a type I square over a type II square, with any subsequent ties being resolved by an equal probability random choice from the remaining candidate squares.

We plot a distribution generated by this model in Fig. 3.11. We find this distribution to be multifractal, with a $D_q$ that is similar in shape to, but everywhere larger than, the $D_q$ of Model 2 (except for $q = 0$ where the two coincide with $D_0 = 2$). In particular, the information dimension is $D_1 = 1.30 \pm 0.02$ (averaged over 20 distributions of $10^6$ points), as compared to $D_1 \simeq 1.05$ for

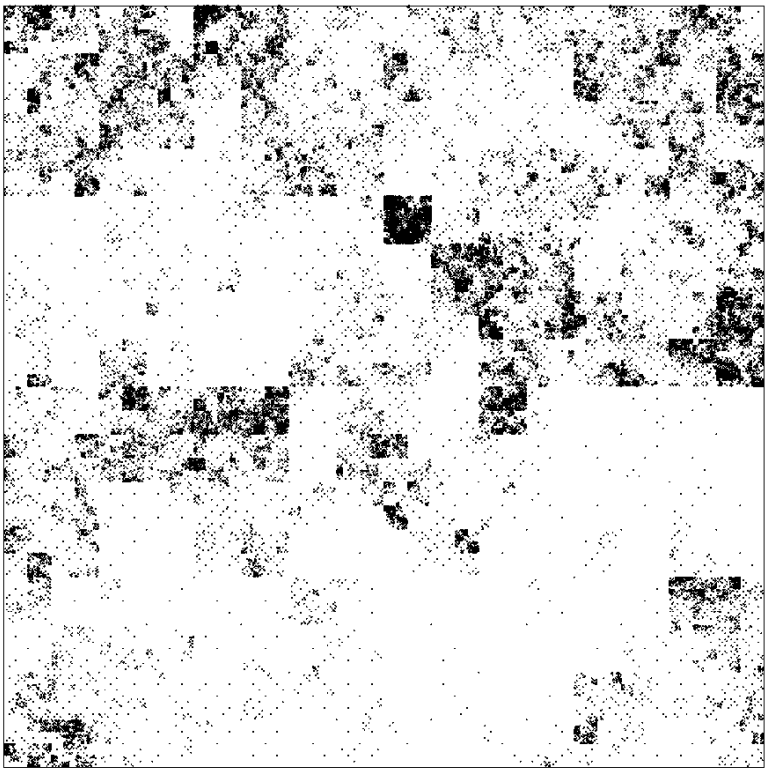

Figure 3.11: A plot of a $10^5$ point distribution on the unit square generated by Model 4.

Model 2 (Eq. 3.29).

### 3.2.5 2-D Unstructured (Model 5)

Looking at Fig. 3.7 (Model 2) and Fig. 3.11 (Model 4), we notice square shaped artifacts and diagonal point formations that arise due to the structured, square-based nature of those point placement schemes. Model 5 is a two dimensional point placement scheme that aims to avoid such artifacts, while maintaining the two ingredients of reproduction and local resettlement. We begin with a point at $(0.5, 0.5)$ and assign to this point a *child distance* $d_c = 1/4$. Step (a) is identical to that used in Models 1-4. Once the target point is selected, Step (b) is to place the new point at a random location on the circumference of a circle centered on the target point and with radius

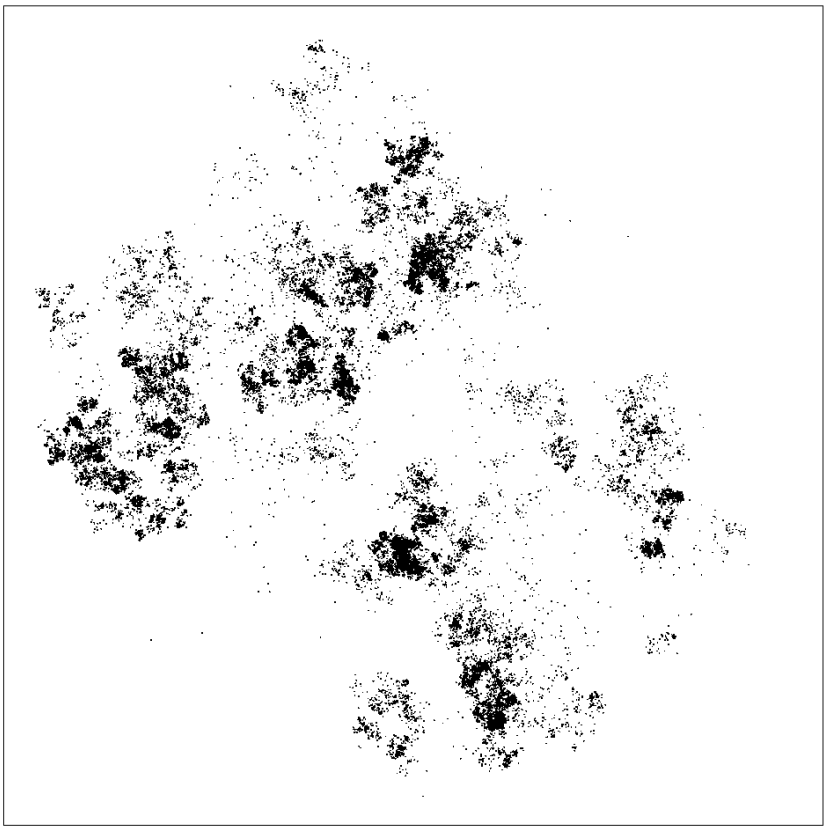

Figure 3.12: A plot of a $10^5$ point distribution on the unit square generated by Model 5.

equal to the target point's $d_c$. A child distance equal to half the target point's $d_c$ is assigned to the new point.

A distribution generated by Model 5 is shown in Fig. 3.12. As intended, there are no grid-type artifacts visible. The distributions generated by this model are found to be multifractal with $D_1 = 1.36 \pm 0.05$ (averaged over 20 distributions of $10^6$ points).

### 3.2.6 2-D Square (Model 6)

In this model we consider points to be located at the *centers* of squares. Beginning with a single point in the center of a unit edge length square, we consider reproduction by a process mimicking

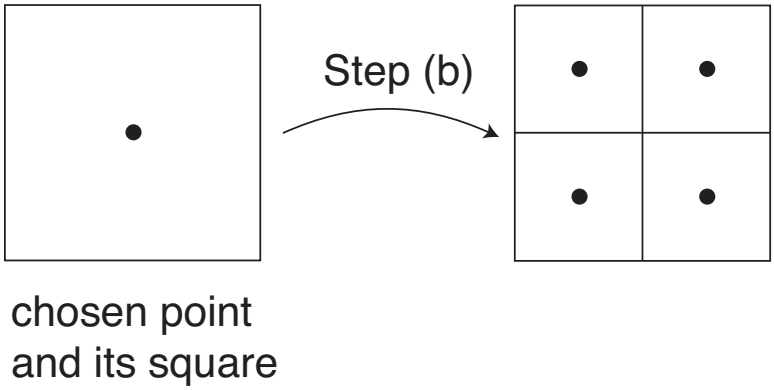


Figure 3.13: Illustration of Step (b) of Model 6.

cell division: In Step (b) a point that was randomly chosen in Step (a) divides into four points (four-fold cell division), all of which then disperse to occupy (resettle) the centers of the four equal size squares obtained by partitioning the original square; see Fig. 3.13.

Note that this construction can be specified without reference to points, but only to the squares they occupy: We can consider Step (a) as randomly choosing a square with equal probability (independent of the size of the square), and Step (b) as dividing the chosen square into four equal new squares. After many applications of this area partitioning procedure, the corresponding population distribution can be produced by simply placing a point in the center of each square.

Figure 3.14 is a plot of the average $D_q$ of 20 independent distributions generated by this model, where the solid line is the analytical result (3.34) derived in Sec. 3.5.3. Both the numerical and analytical results confirm that Model 6 generates multifractal distributions.

### 3.2.7 2-D Triangle (Model 7)

Model 7 is a slight variation of Model 6. Here triangles, as opposed to squares, are the basis of the construction. We begin with a single equilateral triangle with unit edge length. At each time step: (a) a target point (triangle) is chosen from all the preexisting points (triangles) with equal probability; (b) the target point's triangle is divided into four identical equilateral triangles with edge length equal to half the edge length of the original triangle (see Fig. 3.15), with points placed in their centers. We show in Sec. 3.5.3 that Model 7 has the identical analytical expression for $D_q$

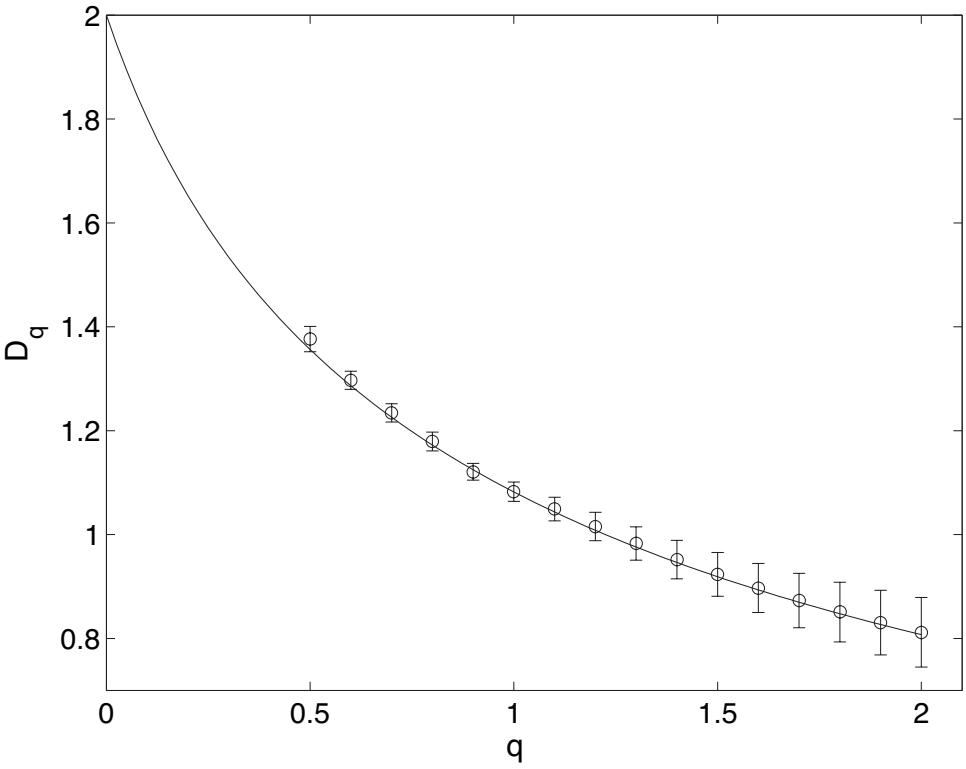


Figure 3.14: Plot of $D_q$ for Model 6. The open circles are the average numerical values of $D_q$ (see caption in Fig. 3.5) obtained from 20 realizations of $10^6$ point distributions generated by this model and the error bars are the sample standard deviations for each value. The solid line is the analytical result of Eq. (3.34).

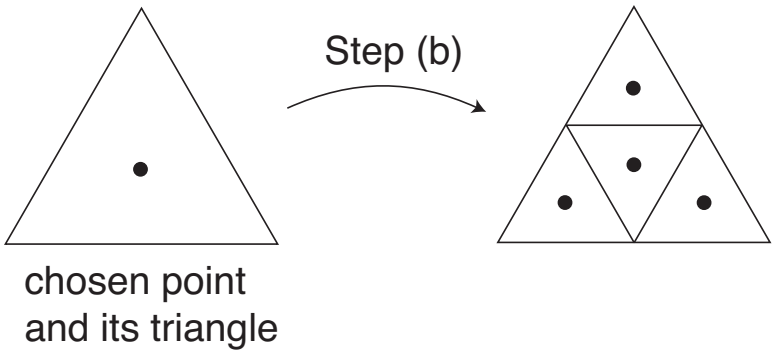


Figure 3.15: Illustration of Step (b) of Model 7.

Table 3.1: A summary of numerical and theoretical results for the point placement models and the Earth at Night image (Fig. 3.1).

| Model | Dimensionality of Space | $D_1$ | $D_q$ (theory) |
|---|---|---|---|
| 1 | 1 | $\frac{1}{2\ln 2} \simeq 0.72$ | Eq. (3.14) |
| 2 | 2 | $\frac{8}{11\ln 2} \simeq 1.05$ | Eq. (3.28) |
| 3 | 1 | $0.84 \pm 0.01$ | - |
| 4 | 2 | $1.30 \pm 0.02$ | - |
| 5 | 2 | $1.36 \pm 0.05$ | - |
| 6,7 | 2 | $\frac{3}{4\ln 2} \simeq 1.08$ | Eq. (3.34) |
| Earth at Night | 2 | $1.60 \pm 0.02$ | - |

as Model 6, and thus also results in multifractal distributions.

Table 3.1 summarizes results for $D_1$ for our models and for the Earth at Night. Table 3.2 summarizes metaphorical interpretations of the reproduction and local resettlement characteristics of our seven models.

Our result that the different strategies of reproduction and local resettlement used in these models (Table 3.2) all lead to multifractal spatial patterns of population density strongly suggests that multifractality may be a generic feature in real situations in which processes involving reproduction and local resettlement take place.

Table 3.2: Summary of reproduction and local resettlement strategies employed in Models 1-7.

| Model | Reproduction | Resettlement |
|---|---|---|
| 1 | parent has one child | child settles randomly chosen adjacent interval |
| 2 | parent has one child | child settles randomly chosen adjacent square |
| 3 | parent has one child | child settles largest adjacent interval |
| 4 | parent has one child | child settles largest and sparsest adjacent square |
| 5 | parent has one child | child settles in a random direction |
| 6 | parent reproduces by four-fold cell division | offsprings share equally the original square area occupied by their parent |
| 7 | parent has three children | parent remains in place; offsprings settle equal shares of parent's triangular area |
| | or | |
| | parent reproduces by four-fold cell division | offsprings share equally the original triangular area occupied by their parent |

We emphasize that the spirit of our approach is very minimalist. In many real cases multiple interacting complex processes undoubtedly influence the determination of population patterns. For example, several likely candidates for the EaN image are geography (mountains, rivers, deserts, etc.), politics, societal and cultural factors, economics, etc. Our models show that considerations of complex spatially heterogeneous processes are not required for explaining the existence of fractally heterogeneous distributions: even very simple dynamics incorporating reproduction and local resettlement are sufficient.

### 3.3 The Effect of Clipped Data on Multifractality

We now revisit the Earth at Night image (Fig. 3.1) and analyze its fractal properties. Figure 3.16(a) is a plot of $D_q$ calculated from the measured light distribution, where the error bars indicate the uncertainty involved in extracting the slopes of the various scaling regions. One feature that is clear is that $D_q$ appears to flatten out (i.e., there is very little variation with $q$) for $q > 0.8$, a characteristic not found in the $D_q$ plots of the models we have examined in Sec. 3.2. We now show how this discrepancy can be resolved.

Figure 3.17 is a log-log plot of the histogram of pixel intensities. A spike is observed for a range of intensity values near the maximum value. This suggests that some regions with high light intensities caused some of the photoelectric cells of the satellites' sensors to saturate, thereby clipping the intensity values at a maximum allowable value. Additionally, it is known that when a photoelectric cell is subjected to high intensities, it can trigger surrounding cells to register more light, resulting in the so called "blooming" effect, which may account for the broadness of the observed spike. We note that the pixels that are part of the spike account for roughly 3% of the total number of non-zero intensity pixels and 10% of the total measured intensity (we consider any pixel with intensity value greater than 209, out of a maximum intensity value of 218, to be part of the spike).

To illustrate the effects that this type of clipping has on multifractal distributions, we apply a scheme that mimics the clipping in the EaN image to a point distribution generated by one of our

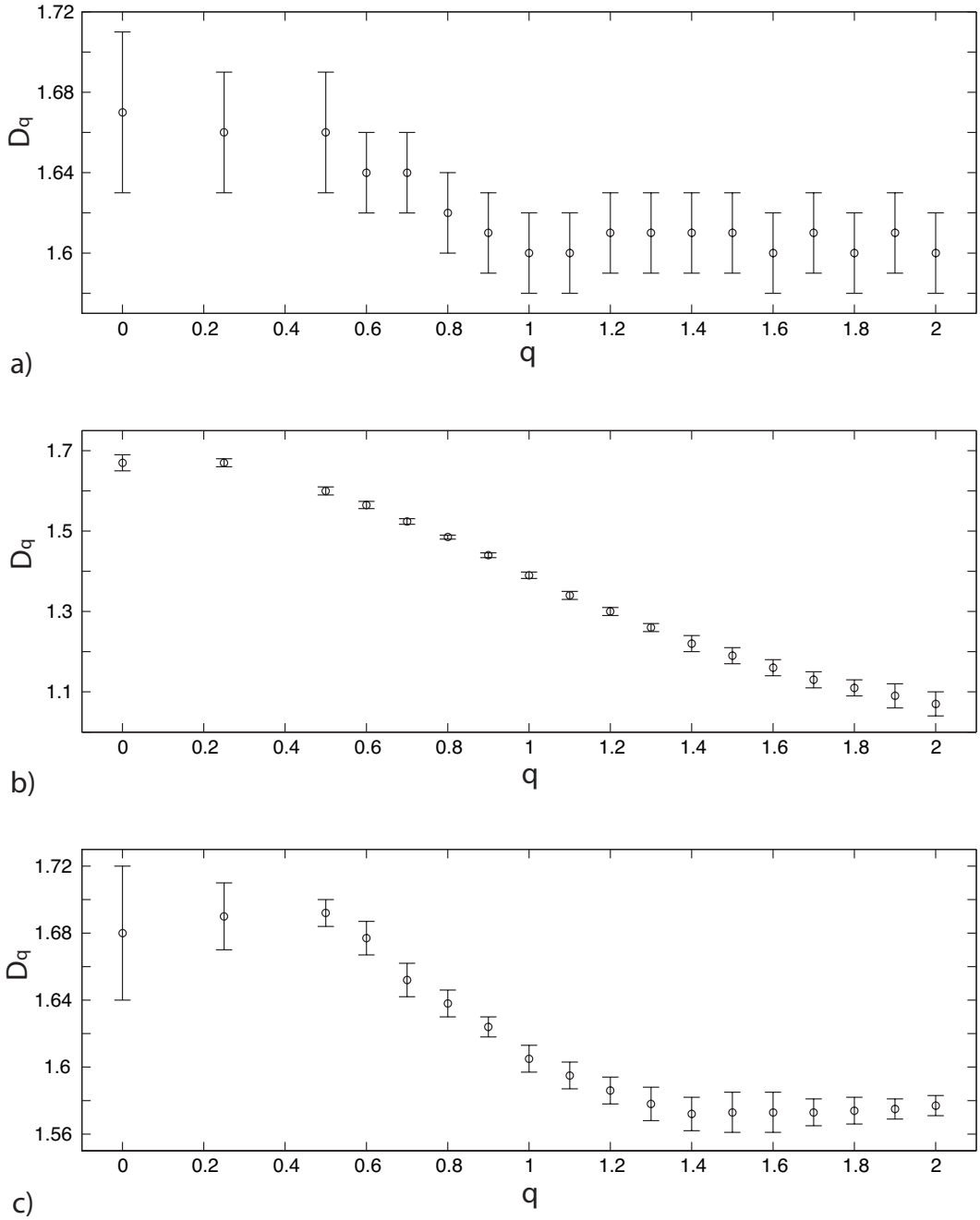


Figure 3.16: a) Plot of $D_q$ for the Earth at Night image. b) Plot of $D_q$ for a $10^6$ point distribution generated by Model 5 before the clipping procedure discussed in Section 3.3 is applied, and c) after it is applied (note the different vertical scales). The error bars reflect the uncertainty involved in determining the slopes of the scaling regions in the plots of the quantity $(1-q)^{-1}\ln I(q,\epsilon)$ versus $\ln(1/\epsilon)$ [see Eq. (3.1)].

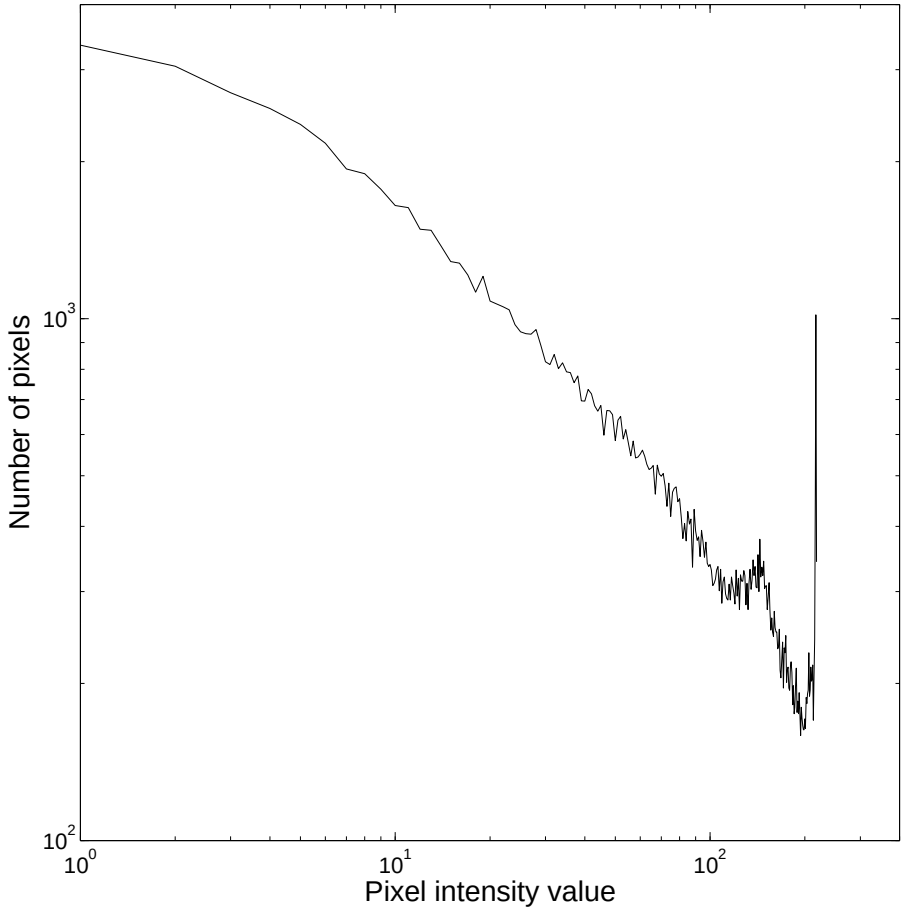


Figure 3.17: Histogram of individual pixel intensities for the Earth at Night image.

models. For this purpose we choose a $10^6$ point distribution generated by Model 5. (Similar results are obtained using our other models.) First we match the smallest scale of our distribution to that of the EaN image. Since the EaN image we analyzed is made up of $2400 \times 800$ pixels, we cover our distribution with a fixed size grid of simulated pixels in which a simulated pixel corresponds to a grid square with edge length of $1/1400 \simeq 1/\sqrt{2400 \times 800}$. Next, we choose the clipping value such that the total number of clipped simulated pixels make up 3% of the total nonempty simulated pixels. For this particular distribution, any simulated pixel containing more than 58 points has its count reset to 58.

Figures 3.18(a) and 3.18(b) are histograms showing the distribution of the number of points contained in individual simulated pixels before and after the clipping is applied, respectively. We see that the clipped histogram is qualitatively similar to the histogram for the EaN image (Fig. 3.17). Next, we calculate $D_q$ for the point distribution before and after clipping (Fig. 3.16(b) and Fig. 3.16(c), respectively). Noting the different vertical scales, we see that the clipping procedure transforms the $D_q$ of Fig. 3.16(b) to a more gradual curve [Fig. 3.16(c)] with an apparent constant tail for the larger $q$ values, making it similar to the $D_q$ for the EaN image [Fig. 3.16(a)]. Additionally, we note that the value of $D_1$ is raised by the clipping (to approximately 1.61 from approximately 1.39). Comparing Fig. 3.16(a) and Fig. 3.16(c), we are able to conclude that the EaN image characteristics are consistent with a multifractal light intensity distribution sensed by an instrument that saturates at a maximum measurable intensity value.

## 3.4 Inhomogeneity

Our models can be generalized to include geographical inhomogeneity. For example, one can imagine that the underlying space is supplemented by a fertility field, and that reproduction is more likely in regions of higher fertility. To investigate the effects of such inhomogeneities on the $D_q$ of a point distribution, we modify Step (a) of Model 6 by use of a fertility field $F(x, y)$ which we use to construct a space-dependent, parent node selection probability. If the location of node $i$ is $(x_i, y_i)$, then the selection probability for this node is taken to be $P_i = F(x_i, y_i)/\sum_j F(x_j, y_j)$.

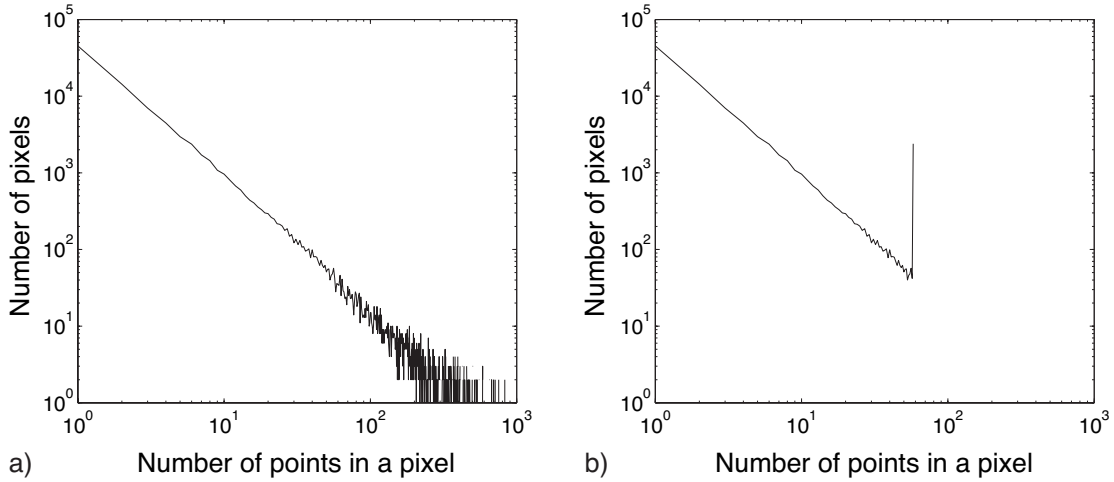


Figure 3.18: Histograms of the number of points in individual simulated pixels (boxes in an $\epsilon = 1/1400$ grid) used to cover a $10^6$ point distribution generated by Model 5, a) for the original distribution and b) after the original distribution was clipped by allowing a maximum of 58 points in any simulated pixel.

We investigate two forms for the fertility field. First we use a smooth field

$$F_s(x, y) = 1 - 0.3\,\{\cos(2\pi x) + \sin(2\pi y)\} \tag{3.4}$$

on the unit square ($0 \le x \le 1, 0 \le y \le 1$). We generate point distributions with this modified version of Model 6 and find that, while the regions of larger $F_s$ in the center of the square are much more dense with points than the regions of smaller $F_s$ near the corners, there is, nevertheless, no discernible difference in $D_q$ compared to the homogeneous case. That is, both the values of $D_q$, as well as the quality and extent of the scaling ranges, remain unchanged.

Next, to explore whether the above result can be attributed to the local smoothness of $F_s$, we consider a rough fertility field,

$$F_r(x, y) = 1 + 0.1\,[f_w(x) + f_w(y)]\,, \tag{3.5}$$

$$f_w(z) \equiv -\sum_{j=0}^{\infty} \alpha^{-j} \cos(2\pi\beta^j z). \tag{3.6}$$

with $\alpha = 1.5$ and $\beta = 3$. For $\alpha < \beta$, the function $f_w(z)$ is a 'Weierstrass function' (see Fig. 3.19); it is rough in the sense that, although it is continuous, it is nondifferentiable, and the graph of $f_w(z)$

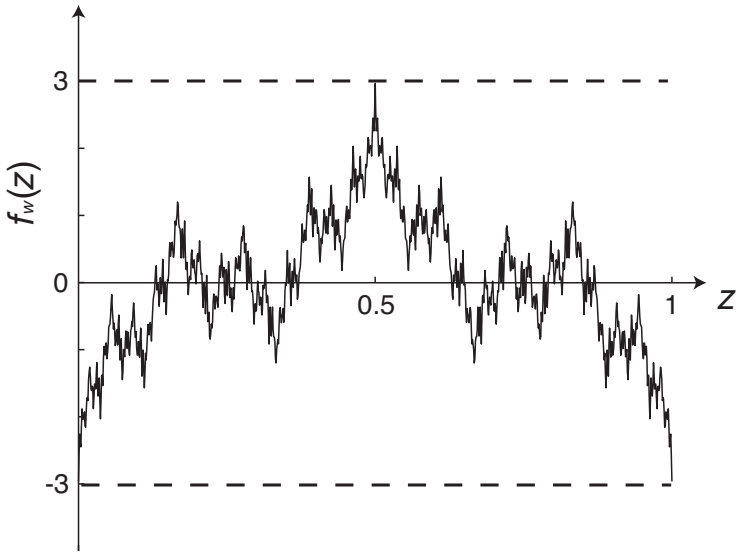


Figure 3.19: A plot of the Weierstrass function $f_w(z)$ [Eq. (3.6)]. In our numerical implementation the infinite upper limit on the summation in Eq. (3.6) is replaced by $j_{max}$, where $j_{max}$ is chosen such that the smallest value of $\epsilon$ used in determining $D_q$ is at least one order of magnitude larger than the smallest scale of the roughness, $3^{-j_{max}}$.

versus $z$ is a fractal curve (the fractal dimension of the curve is $2-(\ln\alpha)/(\ln\beta)\simeq 1.63$ [24]). With $\alpha = 1.5$, $F_r$ has the same range of variation as $F_s$ (i.e., 0.4 to 1.6). Applying this fertility field to Model 6 results in the same values and scaling ranges for $D_q$ as in the homogeneous case and the smooth fertility field case [Eq. (3.4)]. This suggests that, while inhomogeneities of the underlying space may dictate certain aspects of the distribution of a growing population, for example, where the points are more likely to settle, its fractality can be due mainly to reproduction and local resettlement processes.

## 3.5 Theory

### 3.5.1 Model 1

Here we derive an analytical expression for the fractal dimension spectrum $D_q$ of population distributions generated by the one dimensional point placement scheme of Model 1. We begin by making the following observation. Choosing a target interval by first selecting a target point from all the preexisting points with equal probability and then selecting between the two adjacent intervals of the target point with equal probability is equivalent to choosing a target interval out of all the preexisting intervals with equal probability, up to the negligible effect (for large time) of the edge intervals that are only bounded on one of their sides by a point. (Alternatively, one could impose periodic boundary conditions on the unit interval, which would result in complete equivalence of the two constructions.)

Having made this identification, we define $N(k,t)$ to be the average (over different realizations of Model 1) number of intervals of length $2^{-k}$ at time $t$, where at every integer value of $t$ a new point is added. The initial condition for this model (at $t = 1$) is a single point at $x = 0.5$ on the unit interval ($0 \leq x \leq 1$). This defines two intervals of length $1/2$ and thus $N(k,1) = 2\delta_{k1}$, where $\delta_{kj}$ is the Kronecker delta. At each time $t$, the probability that an interval of length $2^{-k}$ is chosen is $N(k,t)/\sum_j N(j,t)$. Then, in place of such a chosen interval, two intervals of length $2^{-(k+1)}$ are created. Since we begin with two intervals at $t = 1$ and at each subsequent time step an interval is destroyed while two are newly created, we see that $\sum_j N(j,t) = t + 1$. The discrete time evolution equation for $N$ is then

$$N(k,t+1) = \left(1 - \frac{1}{t+1}\right) N(k,t) + \frac{2}{t+1} N(k-1,t), \tag{3.7}$$

where $1/(t+1)$ is the probability that a given interval is selected at time $t$, $(t+1)^{-1}N(k,t)$ is the average (over different realizations of Model 1) number of intervals of length $2^{-k}$ destroyed at time $t+1$, and $2(t+1)^{-1}N(k-1,t)$ is the average number of $2^{-k}$ length intervals created at time $t+1$. Alternatively, in terms of $n(k,t)$, the fraction of the total number of intervals at time $t$ with

length $2^{-k}$, given by $n(k,t) = N(k,t)/(t+1)$, the evolution equation is

$$n(k,t+1) = \left(\frac{t}{t+2}\right) n(k,t) + \left(\frac{2}{t+2}\right) n(k-1,t). \tag{3.8}$$

In Appendix C, we introduce the partition function formalism for calculating $D_q$. We define the quantity $\tau = (q-1)D_q$ and derive the following expression for $q(\tau)$ for Model 1 [Eq. (C.10)]:

$$q(\tau) = 1 + \lim_{t\to\infty} \frac{\ln\langle 2^{k\tau}\rangle_t}{\ln t}, \quad \langle 2^{k\tau}\rangle_t = \sum_k n(k,t) 2^{k\tau}. \tag{3.9}$$

Multiplying both sides of Eq. (3.8) by $2^{k\tau}$ and then summing over $k$, we obtain the recursion relation,

$$\langle 2^{k\tau}\rangle_{t+1} = \langle 2^{k\tau}\rangle_t \left(1 + \frac{2^{\tau+1}-2}{t+2}\right), \tag{3.10}$$

which, since $\langle 2^{k\tau}\rangle_{t=1} = 2^\tau$, can be solve to yield,

$$\langle 2^{k\tau}\rangle_t = \prod_{i=1}^{t} \left(1 + \frac{2^{\tau+1}-2}{i+1}\right). \tag{3.11}$$

Substituting Eq. (3.11) into Eq. (3.9), we obtain

$$q = 2^{\tau+1} - 1, \tag{3.12}$$

where we have made use of the relationship

$$\lim_{t\to\infty} \frac{\sum_{i=a}^{t} \ln\left(1 + \frac{x}{i+y}\right)}{\ln t} = x, \tag{3.13}$$

for $x$ and $y$ constants and $a$ an integer such that $a + y + x > 0$ and $a + y > 0$. Then, since $D_q = \tau(q)/(q-1)$, we invert Eq. (3.12) for $\tau(q)$ and obtain the following result for $D_q$:

$$D_q = \frac{\log_2(q+1) - 1}{q-1}. \tag{3.14}$$

In particular, using L'Hospital's Rule,

$$D_1 = \frac{1}{2\ln 2} \simeq 0.72. \tag{3.15}$$

### 3.5.2 Model 2

The calculation of $D_q$ for Model 2 is based on the classification of squares into types I and II, discussed in Sec. 3.2.4. We begin by defining $N_1(k,t)$ and $N_2(k,t)$ to be the average (over different

realizations) number of type I and type II squares, respectively, with edge length $2^{-k}$ at time $t$. The initial state ($t = 1$) is a single point at $(0.5, 0.5)$. The point divides the unit square into four equal squares of edge length $1/2$, with each square having one point on one of its vertices. Thus $N_1(k, 1) = 4\delta_{k1}$, while $N_2(k, 1) = 0$ for all $k$. At each time $t$, the new point is equally likely to appear in each type I square, but is twice as likely to appear in each type II square, because there are two parent points that can produce offspring in each type II square. In other words, a type I square with edge length $2^{-k}$ is chosen with probability $N_1(k,t)/\sum_j (N_1(j,t) + 2N_2(j,t))$, while a type II square of the same size is chosen with probability $2N_2(k,t)/\sum_j (N_1(j,t) + 2N_2(j,t))$. When a type I square is chosen, the new point is placed in its center, destroying the original square and creating three type I squares and one type II square, all with half the edge length of the original square. On the other hand, when the new point is in a type II square, it is replaced by two type I squares and two type II squares, all having half the edge length of the original square. Notice that either way, the quantity $\sum_j (N_1(j,t) + 2N_2(j,t))$ increases by 4 between time $t$ and $t + 1$, so that $\sum_j (N_1(j,t) + 2N_2(j,t)) = 4t$.

The discrete time evolution equations for both types of squares are then

$$N_1(k, t+1) = N_1(k,t)\left(1 - \frac{1}{4t}\right) + N_1(k-1,t)\frac{3}{4t} + N_2(k-1,t)\frac{4}{4t}, \tag{3.16}$$

$$N_2(k, t+1) = N_2(k,t)\left(1 - \frac{2}{4t}\right) + N_1(k-1,t)\frac{1}{4t} + N_2(k-1,t)\frac{4}{4t}. \tag{3.17}$$

In Appendix C we show that for Model 2 [Eq. (C.12)],

$$q(\tau) = \lim_{t\to\infty} \frac{\ln\left(T_1(t) + 2^q T_2(t)\right)}{\ln t}, \tag{3.18}$$

where $T_i(t) = \sum_k 2^{k\tau} N_i(k,t)$. Multiplying Eqs. (3.16) and (3.17) by $2^{k\tau}$ and summing over $k$, we obtain

$$T_1(t+1) = T_1(t)\left(1 + \frac{3\cdot 2^\tau - 1}{4t}\right) + T_2(t)\frac{4\cdot 2^\tau}{4t}, \tag{3.19}$$

$$T_2(t+1) = T_2(t)\left(1 + \frac{4\cdot 2^\tau - 2}{4t}\right) + T_1(t)\frac{2^\tau}{4t}. \tag{3.20}$$

At this point, we make the continuous time approximation $T_i(t+1) - T_i(t) \approx \mathrm{d}T_i(t)/\mathrm{d}t$, valid for large $t$, which allows us to represent Eqs. (3.19) and (3.20) as the matrix differential

equation,

$$\frac{\mathrm{d}\mathbf{T}(s)}{\mathrm{d}s} = \mathbf{M} \cdot \mathbf{T}(s), \tag{3.21}$$

where

$$\mathbf{T}(s) = \begin{bmatrix} T_1(s) \\ T_2(s) \end{bmatrix}, \tag{3.22}$$

$$\mathbf{M} = \begin{bmatrix} 3 \cdot 2^\tau - 1 & 4 \cdot 2^\tau \\ 2^\tau & 4 \cdot 2^\tau - 2 \end{bmatrix} \tag{3.23}$$

and $s = (1/4) \ln t$.

Equation (3.21) is solved in terms of the eigenvalues $\lambda_a, \lambda_b$ and the eigenvectors $\mathbf{T}_a, \mathbf{T}_b$ of the matrix $\mathbf{M}$, giving

$$\mathbf{T}_i \sim e^{\lambda_i s} = t^{\frac{\lambda_i}{4}}, \tag{3.24}$$

for $i = a, b$, with $\lambda_a, \lambda_b$ distinct and real and $\lambda_a > \lambda_b$. We can express the quantity $(T_1(t) + 2^q T_2(t))$ in Eq. (3.18) in terms of a linear combination of the components of these eigenvectors, whose long time behavior is dominated by the behavior of $\mathbf{T}_a$, the eigenvector associated with the larger eigenvalue $\lambda_a$. Thus (3.18) gives us the simple result

$$q = \lambda_a / 4. \tag{3.25}$$

The eigenvalues of $\mathbf{M}$ are the roots of its characteristic polynomial,

$$\lambda^2 + \lambda(3 - 7 \cdot 2^\tau) + 2 - 10 \cdot 2^\tau + 8 \cdot 2^{2\tau} = 0. \tag{3.26}$$

This equation is also quadratic in $2^\tau$, allowing us to solve for $\tau(\lambda)$

$$\tau(\lambda) = \log_2 \left[ \frac{1}{16} \left( 10 + 7\lambda \pm \sqrt{17\lambda^2 + 44\lambda + 36} \right) \right]. \tag{3.27}$$

We substitute the relationship from Eq. (3.25) into Eq. (3.27) and, since $\tau(q) = (q-1)D_q$, we resolve the ambiguity of the $\pm$ sign by requiring that $\tau(q = 1) = 0$. This results in

$$D_q = \frac{1}{q-1} \log_2 \left[ \frac{1}{8} \left( 5 + 14q - \sqrt{68q^2 + 44q + 9} \right) \right] \tag{3.28}$$

for the dimension spectrum of Model 2. In particular, applying L'Hospital's Rule to (3.28), the information dimension is

$$D_1 = \frac{8}{11 \ln 2} \simeq 1.05. \tag{3.29}$$

### 3.5.3 Models 6 and 7

Although Model 6 is a two dimensional point placement scheme, the fact that each square is equally likely to be chosen at a given time allows the calculation of $D_q$ to be very similar to that for the one dimensional scheme of Model 1. We begin by defining $N(k,t)$ to be the average number of squares with edge length $2^{-k}$ at time $t$. The initial state (at $t = 1$) is a single square with unit edge length, giving us $N(k,1) = \delta_{k0}$. At each time $t$, a square with edge length $2^{-k}$ is chosen with probability $N(k,t)/\sum_j N(j,t)$ and replaced by four equal squares with edge lengths $2^{-(k+1)}$. Since at each time step there is a net gain of three squares, we see that $\sum_j N(j,t) = 3t - 2$, and obtain the following discrete time evolution equation for $N$:

$$N(k,t+1) = N(k,t)\left(1 - \frac{1}{3t-2}\right) + N(k-1,t)\frac{4}{3t-2}. \tag{3.30}$$

In terms of the fraction $n(k,t) = N(k,t)/(3t-2)$ of squares with edge length $2^{-k}$, we have

$$n(k,t+1) = n(k,t)\frac{3t-3}{3t+1} + n(k-1,t)\frac{4}{3t+1}. \tag{3.31}$$

Interchanging the squares in the above derivation with equilateral triangles, one can see that the same evolution equations hold for the triangles in Model 7.

In Appendix C we show that Eq. (3.9), which allowed us to calculate $q(\tau)$ in terms of $\langle 2^{k\tau}\rangle_t$ for Model 1, holds for Models 6 and 7 as well. Following the procedure in Section 3.5.1, we multiply both sides of Eq. (3.31) by $2^{k\tau}$ and sum over $k$, solving the resulting recursion relation to obtain

$$\langle 2^{k\tau}\rangle_t = \prod_{i=2}^{t}\left(1 + \frac{(2^\tau - 1)\,4/3}{i - 2/3}\right). \tag{3.32}$$

Making use of the relationship in Eq. (3.13), we get

$$q(\tau) = 1 + \frac{4}{3}\left(2^\tau - 1\right). \tag{3.33}$$

This equation is inverted to solve for $\tau(q)$ which, when divided by $(q-1)$, gives us the multifractal dimension spectrum

$$D_q = \frac{\log_2\left((q-1)\frac{3}{4}+1\right)}{q-1}. \tag{3.34}$$

The information dimension is

$$D_1 = \frac{3}{4\log 2} \simeq 1.08. \tag{3.35}$$

## 3.6 Conclusion

The main conclusion of this chapter is that reproduction and local resettlement processes may lead to multifractal spatial distributions for growing populations. We introduced a number of point placement models in one and two dimensions and showed that the models resulted in multifractal distributions. Furthermore, we have demonstrated qualitative similarity between the example of the Earth at Night image and clipped versions of distributions generated by our models. We thus suggest that the mechanism by which our models create multifractal distributions may be operative in the growth of real systems.

Chapter 4

# Simulating the Segregation of Ternary Granular Mixtures in Long Rotating Cylinders

## 4.1 Introduction

An interesting property of granular systems is the tendency of mixtures of different types of materials to unmix (e.g., under vertical vibration [28] or during avalanching [29]). A binary mixture of granular particles with different physical properties (e.g., different friction coefficients, sizes, densities) is known to exhibit segregation in a half filled horizontal cylinder rotating about its axis, a phenomenon of interest to the many areas of industry that use cylindrical rotating drums for various processes such as mixing and grinding. Within the first few rotations the mixture segregates radially, with the less mobile particles making up a core that extends in the axial direction (along the cylinder axis of rotation), while the more mobile particles are pushed out radially to the cylinder walls. Following radial segregation, if the proper conditions are met (rotational frequency, particle properties, etc.), the particles can also segregate axially. When viewing the surface of of the granular mixture one sees axial banding, alternating bands of regions rich in each of the particle types, along the cylinder axis. Over long periods of time, the band structure is observed to coarsen, with adjacent bands merging, until a final stable state is reached consisting of a smaller number of axial bands. This type of segregation was first observed by Oyama in 1939 [30] for an equal volume mixture of glass spheres of two different sizes, and was subsequently studied [31–41] both experimentally and theoretically.

One of the earliest partial successes of understanding the mechanisms responsible for this phenomenon was by Zik et al. [35, 36]. They observed axial segregation in a mixture of sand and glass, particle types that differed in their frictional properties. This led to an analytical model where axial segregation was driven by the difference in the dynamic angles of repose of two different particles types. The dynamic angle of repose for a particle type is the angle of the incline created by the continuously flowing granules. The authors posited that the axial segregation occurred due to a sorting mechanism in the thin flowing layer that formed as the cylinder was rotated, and represented the rotating granular material as a rigidly rotating bulk bottom with a thin top layer of

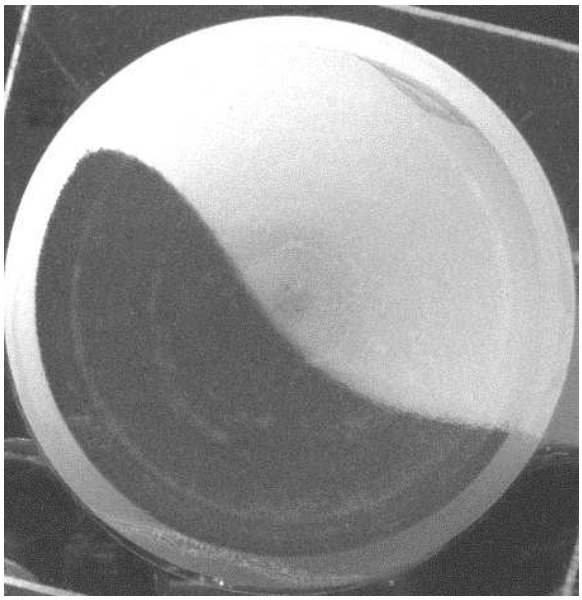

Figure 4.1: The S-shaped profile of a flowing granular mixture in a rotating cylinder.

flowing particles. One of the central results derived from their model was a numerical solution for the profile shape of the granular flow layer which successfully reproduced the characteristic S-shape profile typical of rotating granular material (see Fig. (4.1)). More importantly, they were able to show that instabilities leading to axial segregation could exist within their theoretical framework and gave the conditions for which they would occur.

This approach, however, did not account for the radially segregated core, the precursor to the axial segregation, which was later observed experimentally. Furthermore, using MRI imaging of the bulk of rotating granular mixtures, Hill et al. [33, 34] showed that, even when axial banding is observed at the surface, a core structure existed in the rigidly rotating bulk, a remnant of the earlier radially segregated state. Thus any real understanding of the axial segregation would have to include radial segregation. This observation led to the conjecture that instabilities leading to undulations in the radially segregated core are what eventually cause the axial segregation. Elperin and Vikhansky [38] incorporated this idea into their model for axial segregation. Beginning with an already radially segregated core and allowing for different dynamic angles of repose for the two particle types they considered, they were able to analytically demonstrate the onset of instabilities leading to axial segregation.

In an effort to replicate experimental results, Puri and Hayakawa [40] developed a phenomenological continuum model with which they were able to successfully simulate both the initial radial segregation and the subsequent axial banding, as well as the long term coarsening of the axial bands. The ability to reproduce the experimental results, however, did not shed light on the microscopic processes that enabled the segregation to take place. Yanagita [41] was also able to reproduce a similarly wide range of phenomena. He used a discrete three dimensional lattice model where each granule occupied an individual lattice site. In a manner similar to [35], the rotating mixture of granules was modeled as a rigidly moving bulk with a thin top flowing layer. The favorable feature of Yanagita's model is that the selection rule used to determine the motion of the granules in the flowing layer is simple and relies only on a set of particle-species-dependent generalized friction parameters. This gave rise to the idea that the difference in the flow properties of the different particle types could play a central role in axial segregation.

Recently, Newey et al. [43] carried out experiments involving ternary mixtures of different size glass particles in rotating cylinders. They found that their mixtures segregated into "bands within bands" (see Fig. (4.2)). After the mixture segregated radially, with the largest particles on the outside, followed by the medium sized ones and then the smallest ones making up the innermost core, the large and medium sized particles formed axial bands, in a manner similar to the binary mixture case. Then the small particles formed bands only within the medium particle bands. This phenomenon pointed to, once again, an intimate relationship between the radial and the axial segregation since the ordering of the axial bands mimicked the earlier radially segregated core.

In an effort to understand these "bands within bands" and to further explore the role that the radially segregated core plays in the eventual axial segregation, we developed a discrete three dimensional simulation model, generalizing Yanagita's model of binary mixture segregation [41] to the case of three particle species. Our simulations duplicated many of the general features observed in the ternary mixture experiment [43]. More importantly, the simplicity of our model allowed us to isolate the surface flow properties of the various particle types as playing important roles in this

Figure 4.2: Picture of the "bands within bands" structures observed by Newey et al. [43]. The light blue particles are the largest, the dark blue particles are the intermediate sized ones and the green particles are the smallest.

particular type of segregation phenomena.

In Sec. 4.2 we describe our model. We present data from simulation runs along with discussion of the important features of the results in Sec. 4.3. Section 4.4 summarizes our findings.

## 4.2 The Model

The model, as mentioned above, generalizes a binary mixture model by Yanagita [41] to include three types of particles. We employ a three dimensional square lattice, with discrete integer valued coordinates $(x, y, z)$, and load it with a mixture of three types of particles, where each particle type occupies exactly one lattice site. Here $x$ corresponds to the horizontal radial direction, $y$ is the axial direction (where we impose periodic boundary conditions), and $z$ is the vertical direction. Unlike the experiment of Newey et al. [43], where the particle types differed in size, the model uses particles with different simulated frictional properties to simulate the differences in the surface flows of the particles. Each site in the lattice is assigned a value $\sigma(x, y, z)$ according to the particle type that populates that site, given by

$$\sigma(x, y, z) = \{0, 1, 2, 3\} = \{\text{no particle},\ \text{particle A},\ \text{particle B},\ \text{particle C}\}. \tag{4.1}$$

The lattice is $N_x$ sites in the $x$ direction ($x = 0, \ldots N_x - 1$), $N_y$ sites in the $y$ direction ($y = 0, \ldots N_y - 1$), and $2N_z$ in the $z$ direction ($z = 0, \ldots 2N_z - 1$). In general the simulations are started by placing $N_x \times N_y \times N_z$ particles in the lower half of the $2N_xN_yN_z$ size lattice, in analogy

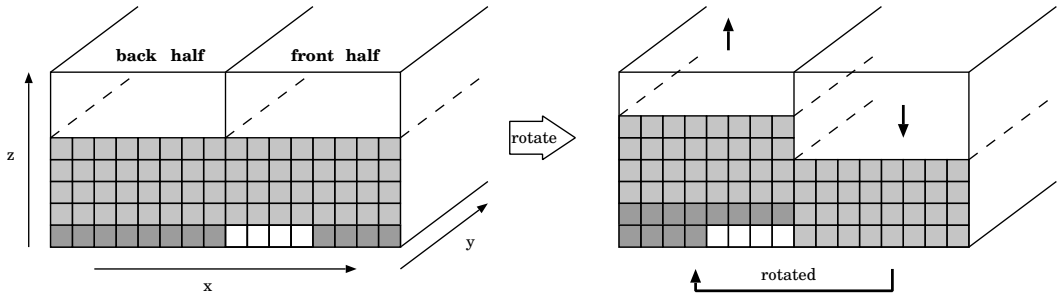


Figure 4.3: Schematic diagram of the discrete rotation process. The lattice is separated into its back ($x < N_x/2$) and front ($x \geq N_x/2$) halves. The back half is raised one step and the bottom slice of the front half is rotated and attached to the bottom of the back half. The front half is then lowered one step.

with a half filled cylinder. The initial states used are usually randomly generated homogeneous mixtures of particles A, B, and C with varying relative abundances, subject to a simulated gravity acting in the negative $z$ direction (i.e. no empty lattice site below any given particle).

The simulations carry out a large ($> 10^5$) number of discrete time steps. Each time step involves two ingredients to mimic (a) the rotation of the cylinder and (b) a flow of the top granular layer:

(a) We first create a height differential between the back ($x < N_x/2$) and the front ($x \geq N_x/2$) of the lattice. This is done by raising the back half of the lattice one step in the positive $z$ direction ($z \to z+1$ for $x < N_x/2$). The bottom slice of the front half of the lattice ($z = 0$, $x \geq N_x/2$) is then reflected about the $y$-axis ($x \to N_x - x - 1$) and placed below the bottom most slice of the back half. The front half is then lowered one step in the negative $z$ direction ($z \to z-1$ for $z \geq 1$ and $x \geq N_x/2$). See Fig. 4.3. This step is meant to mimic rotation in the experimental cylindrical geometry, but adapted to our rectangular simulation geometry.

(b) A particle on the surface $h(x,y) = \sum_{z=0}^{N_z-1} \Theta(\sigma(x,y,z))$ is chosen at random [here $\Theta(\sigma) = 0$ if $\sigma = 0$ (the site is unoccupied) and $\Theta(\sigma) = 1$ if $\sigma > 0$ (the site is occupied)]. This particle will be allowed to move from its position $(x, y, h(x,y)-1)$ to a position $(x+1, y+\Delta y, h(x+1, y+\Delta y))$, with $\Delta y$ randomly chosen from the values $\{-1, 0, 1\}$ (to simulate diffusion in the $y$ direction),

subject to a selection rule. The selection rule compares the difference in height at points $(x, y)$ and $(x+1, y+\Delta y)$, namely $h(x, y) - h(x+1, y+\Delta y)$, to a generalized friction $F(x, y, h(x, y) - 1)$, or resistance to motion in the positive $x$ direction, that the particle experiences at its initial point.[1] If the height difference is large enough to overcome this generalized friction then the particle is allowed to make the move. We note that, since the particles that are at $x = N_x - 1$ are already at the lattice edge, they are not allowed to flow in the $x$ direction, and only move via the discrete rotation mechanism [Step (a)].

The generalized friction for a particle is given by:

$$F(x, y, z) = \sum_{(\delta x, \delta y, \delta z)} f(\sigma(x, y, z)\ \sigma(x + \delta x, y + \delta y, z + \delta z)) \tag{4.2}$$

where the values of $\delta x, \delta y$ and $\delta z$ are chosen such that only the four "touching" lattice sites adjacent to the site $(x, y, z)$ are included in the summation (see Fig. 4.4); i.e.,

$$(\delta x, \delta y, \delta z) = \{(0, 1, 0), (0, -1, 0), (-1, 0, 0), (0, 0, -1)\}. \tag{4.3}$$

Here $f$ is the mapping

$$f : \{0, 1, 2, 3, 4, 6, 9\} \rightarrow \{0, F_{AA}, F_{AB}, F_{AC}, F_{BB}, F_{BC}, F_{CC}\}, \tag{4.4}$$

where $F_{ij}$ is the friction parameter value between particle type $i$ and particle type $j$. In our simulations we set

$$\begin{aligned} F_{CC} &> F_{BB} > F_{AA} \\ F_{BB} &> F_{AB} > F_{AA} \\ F_{CC} &> F_{BC} > F_{BB} \\ F_{CC} &> F_{AC} > F_{AA} \end{aligned} \tag{4.5}$$

The first relation in (4.5) sets the C particles as those which experience the largest resistance to motion in the positive $x$ direction, followed by B particles and then A particles. The final three

[1] This resistance to motion can be thought of as resulting from the actual frictional properties of the particles, or from other particle retarding effects. One example of such a retarding effect occurs in the case of particles of different sizes. Smaller particles are more prone than larger particles to fall through or be caught in interstitial spaces, making their progress down the inclined slope more difficult.

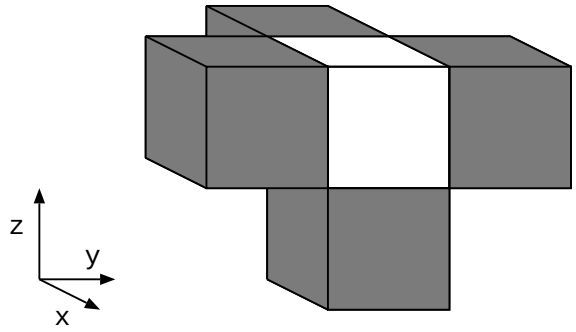


Figure 4.4: An illustration of the adjacent lattice sites, represented by gray cubes, which are included in the summation for calculating the generalized friction $F$ (see Eq. (4.2)) experienced by a particle (white cube).

relations in (4.5) require the friction between different particle types to be some value between their respective "pure" values, as one would expect on physical grounds. The actual values of these generalized friction parameters are adjusted to obtain simulation outcomes most closely resembling experimental results.

The application of Step (b) is carried out $R$ times before the lattice is discretely rotated again (Step (a)). The inverse of $R$ is then analogous to the angular frequency of a rotating cylinder, $\omega \sim 1/R$.

## 4.3 Simulation Results and Discussion

For our simulations we use a lattice with dimensions of 200 lattice sites in the axial direction ($N_y = 200$), 80 lattice sites in the vertical direction ($N_z = 40$), and 10 lattice sites in the horizontal direction ($N_x = 10$), resulting in $1.6 \times 10^5$ total lattice sites. We load the lattice with $8 \times 10^4$ particles, corresponding to a half-filled cylinder, and the relative abundances of the particle types are varied as needed.

Here we present results from simulations run with the following friction parameter values: $F_{AA} = 0.4$, $F_{AB} = 0.44$, $F_{BB} = 1.0$, $F_{AC} = 0.9$, $F_{BC} = 1.5$, and $F_{CC} = 2.4$. While these particular values were found to be optimal, in the sense that, of the friction parameters we tried,

they produced simulation results most closely resembling the experiment, the same qualitative behavior was also seen for a wide range of parameter values, e.g., within a range of at least plus or minus 10% of the chosen values. We found that setting the value of $R$, the number of times that Step (b) is executed in each discrete time step, to $R = 2.5 \times 10^4$ gave good results. Since our lattice has $N_y \times N_x = 2000$ surface particles, this implies that each surface particle experienced 12.5 potential moves on average per rotation time step.

The simulations successfully reproduced a wide range of the experimentally observed phenomena. Beginning with a uniformly mixed state (with equal proportions of each particle type), Fig. 4.5(a) shows the lattice after $2 \times 10^5$ time steps. Particles A are colored black, particles B are yellow and particles C are red. The axial end of the lattice shows evidence of an ordered radial core while the surface layer appears to suggest some form of axial segregation. It is important to note here that the view of the lattice surface is analogous to taking cross sections of granular mixtures in physical cylinders. This is because the lattice is viewed right after we execute $R$ applications of Step (b) (i.e., the particle flow step), and right before the discrete rotation, Step (a), is made. Thus, unlike the experimental situation, there is no flowing top layer obscuring the underlying mixture. Fig. 4.5(b) shows the profile of the mixture surface for various axial ($y$) positions. The profiles qualitatively reproduce the experimentally observed granular flow profiles (cf. Fig. 4.1). As is expected, regions exhibiting a high concentration of C particles, the particle type associated with the largest friction parameters, also have the steepest slopes.

The view of our lattice most closely resembling a laboratory view of the rotating granular mixture is the view of the rear slice (i.e., $x = 0$), shown in Fig. 4.6(a). The "bands within bands" structures observed in [43] are clearly visible, with the red particles mainly found within bands of yellow particles, which are themselves mostly within black particle regions. An advantage of our simulation over the experiment is the fact that we have access to the full three dimensional structure of the granular mixture. This allows us to determine the relative abundance $\phi$ of each particle type in each two dimensional axial slice and we plot the result in Fig. 4.6(b). The interesting feature here is that, while particles A and C vary in concentration between regions of great abundance

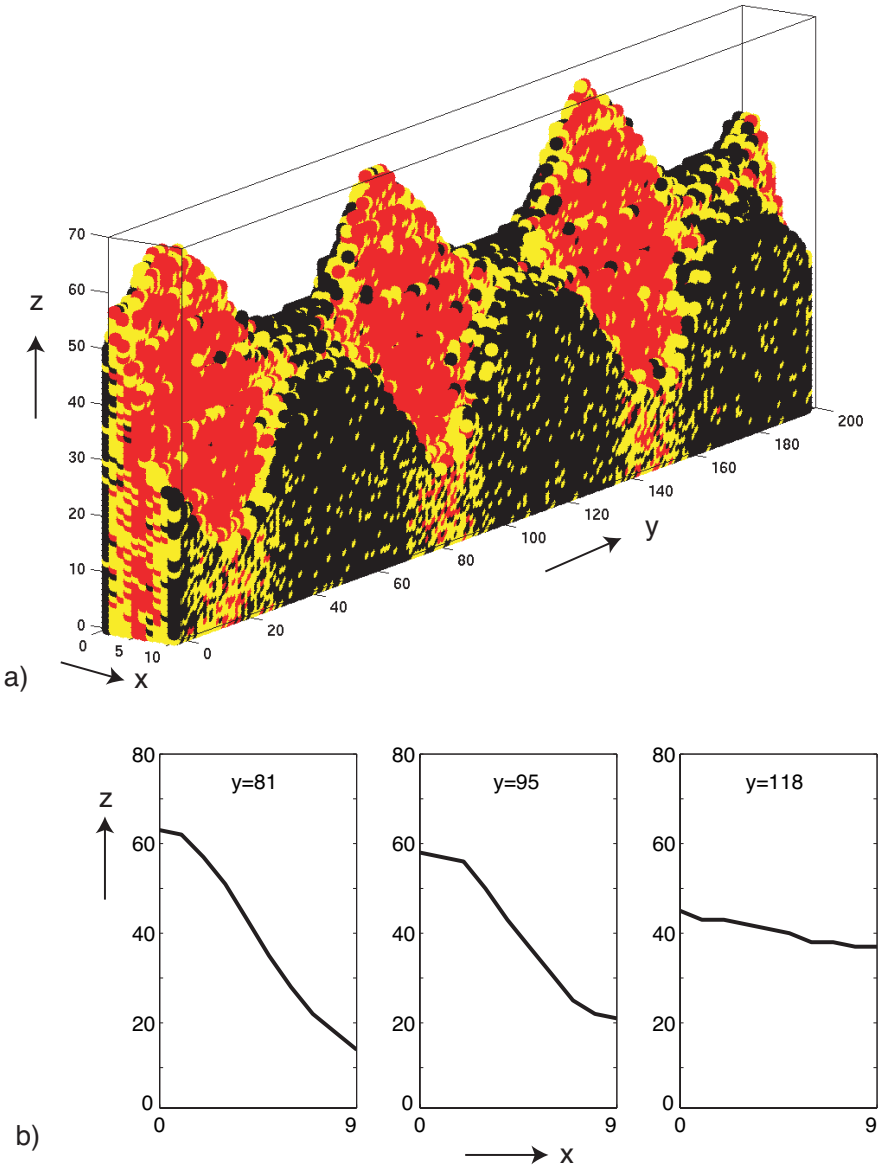


Figure 4.5: a) View of the three dimensional lattice after $2 \times 10^5$ time steps. Initially the lattice is loaded with a uniformly mixed state of particles A, B and C (colored red, yellow and black, respectively), with equal proportions. b) Profiles of the particle mixture surface at three different axial ($y$) positions. [$F_{AA} = 0.4$, $F_{AB} = 0.44$, $F_{BB} = 1.0$, $F_{AC} = 0.9$, $F_{BC} = 1.5$, $F_{CC} = 2.4$, $R = 2.5 \times 10^4$]

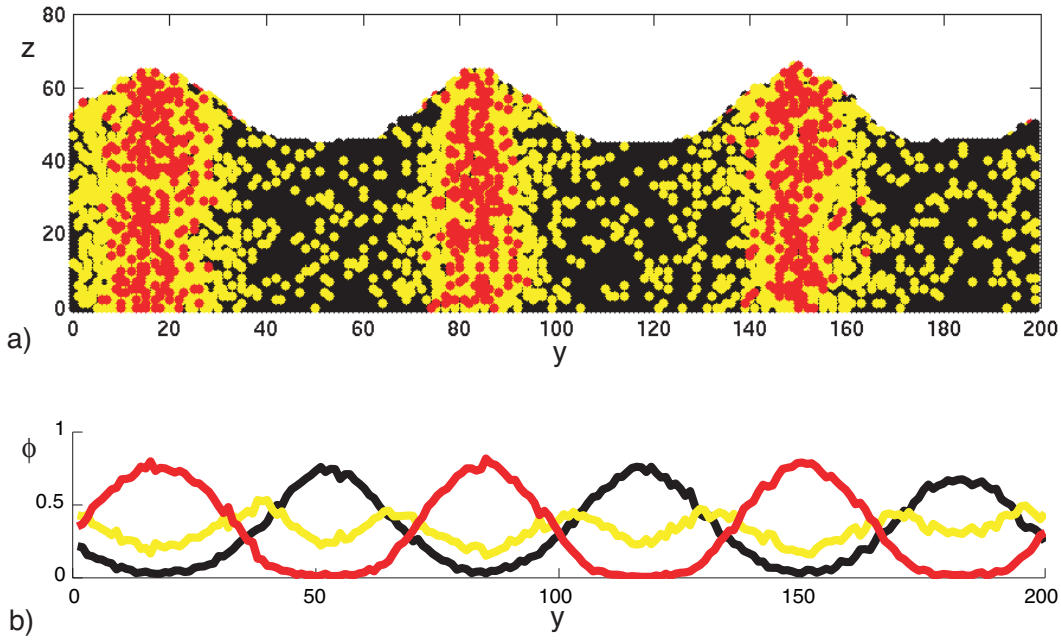


Figure 4.6: a) View from the rear ($x = 0$) of the lattice in Fig. 4.5(a), showing the "bands within bands" structures. This perspective most closely resembles the laboratory view of the surface of a rotating granular mixture (cf. Fig. 4.2). b) The relative abundance $\phi$ of each of the particle types for two dimensional slices ($0 \leq x < N_x$ and $0 \leq z < 2N_z$) at each axial position $y$. [Black signifies particles A, yellow signifies particles B and red signifies particles C.]

and extreme scarcity, particles B are found throughout the axial direction, becoming the dominant particle type only between regions rich in particles A and those rich in particles C.

Fig. 4.7(a) is a space-time plot of a simulation run with 25% A particles, 30% B particles and 45% C particles. In each time step, the rear slice of the lattice is vertically averaged such that the most abundant particle type at each axial position $y$ (the vertical column defined by $x = 0$ and $0 \leq z < 2N_z$) is assigned to that axial position. Here the black areas correspond to A particles, the yellow areas to B particles, and the red areas to C particles. The initial formation of yellow bands followed by the appearance of red bands exclusively within the yellow regions is clearly observed. One can also see the coarsening of the bands via merging.

Figures 4.7(b-e) show top views of the lattice at four different time steps where the arrow from each image indicates the corresponding time in the space-time plot, Fig. 4.7(a). [(b) $t = 0$

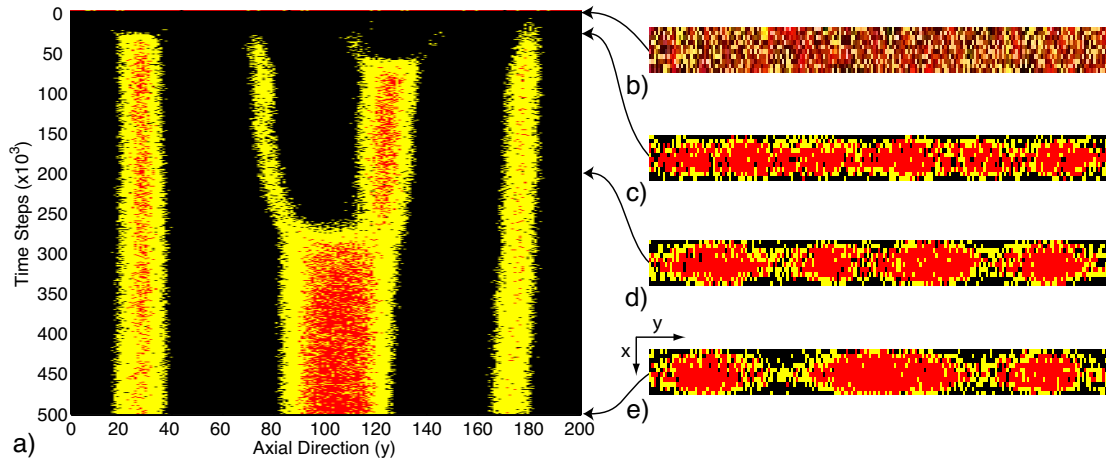


Figure 4.7: a) Spacetime plot of the our simulation for a total of $5 \times 10^5$ timesteps. In each time step, the most abundant particle type in the rear ($x = 0$) slice of the lattice is picked for each axial ($y$) position. The black, yellow, and red regions correspond to particles A, B, and C respectively. The relative abundances of the particles are 25% A particles, 30% B particles and 45% C particles. b)-e) Top views of the lattice corresponding to different times in the space-time plot. [(b) $t = 0$ (c) $t = 2 \times 10^4$, (d) $t = 2 \times 10^5$, (e) $t = 5 \times 10^5$]

(c) $t = 2\times10^4$, (d) $t = 2\times10^5$, (e) $t = 5\times10^5$] As was indicated earlier, these images can be regarded as corresponding to the axial cross sections of an actual rotating granular mixture. The cross sections reveal (b) an initial homogeneous mixed state followed by (c) a radially segregated state, with the low friction particles A forming the outer layer, followed by particles B, and then particles C forming the innermost core. There are visible undulations in the core structure; however, Fig. 4.7(a) indicates that there is no discernible axial banding at this point. By the time of Fig. 4.7(d) the instabilities of the core grow, with the $C$ particles pushing the $B$ particles to the lattice walls ($x = 0$, $x = N_x - 1$) in certain locations. In the space-time plot this manifests itself as mostly yellow bands in a black background with a minimal amount of red regions seen within the yellow bands. Finally, Fig. 4.7(e) reveals a large central bulge of C particles, the result of the two central bands seen in Fig. 4.7(d) merging. The C particles are able to push more of the B particles aside, resulting in the clearly visible red band within the central yellow band in the space-time plot. The axial core structure, consisting of C and B particles, persists throughout the lattice. A similar structure was observed in the rotating drum experiments reported in Ref. [43] through the use of a partially transparent sample and back illumination. We also note that band merging, as occurs in the transition between Fig. 4.7(d) and Fig. 4.7(e), is also observed in the experiment of Ref. [43].

The bulk motion Step a), treats all particles the same, and differing frictional properties of the model particles only matter in the surface flow Step b). Thus, in the model, all segregation, both radial or axial, is due to differing surface flow properties of the three particle types, suggesting that differences in the surface flow properties of the particles in the physical system can play an important role in the experimentally observed segregation.

## 4.4 Conclusion

We have extended the model used by Yanagita [41] for binary granular mixture segregation in a rotating cylinder to allow for three particle species. This was done in an effort to model the experiment of Newey, et al. [43] who observed an axial segregation structure consisting of "bands within

bands" for ternary granular mixtures. The simulation indicates that a model with even a thin flowing layer of particles can exhibit the radial segregation and subsequent axial segregation that were experimentally observed, and that differences in frictional properties suffice to qualitatively model the segregation patterns and the sequences in which they appear. That such a simple simulation, paramaterized only by friction, can accurately reproduce the experimental results shows the basic robustness of the physical phenomena.

Appendix A

Calculation of $\bar{H}(k)$

We substitute $H(k,N) = G(k,N)/N$, the fraction of nodes with degree $k$ at time $N$, into Eq. (2.1), obtaining,

$$(N+1)H(k,N+1) = (N-m)H(k,N) + mH(k-1,N) + \delta_{km}. \tag{A.1}$$

We define $\bar{H}(k)$ to be the $N$ independent solution to (A.1). Substituting into (A.1) and rearranging terms we get

$$\bar{H}(k) = \frac{m}{m+1}\bar{H}(k-1) + \frac{\delta_{km}}{m+1}. \tag{A.2}$$

This recursion relation is solved to yield

$$\bar{H}(k) = \frac{1}{m+1}\left(\frac{m}{m+1}\right)^{k-m} \tag{A.3}$$

for $k \geq m$ and $\bar{H}(k) = 0$ for $k < m$.

Now we show that as $N \to \infty$, $H(k,N)$ approaches $\bar{H}(k)$. Dividing (A.1) by $N+1$ and subtracting from this the same equation but with $\bar{H}(k)$ inserted, we obtain after some algebra

$$H(k,N+1) - \bar{H}(k) = \frac{N-m}{N+1}\left[H(k,N) - \bar{H}(k)\right] + \frac{m}{N+1}\left[H(k-1,N) - \bar{H}(k-1)\right]. \tag{A.4}$$

Letting $F(k,N) = \binom{N}{m+1}\left[H(k,N) - \bar{H}(k)\right]$, we find that (A.4) is equivalent to

$$F(k,N+1) = F(k,N) + \frac{m}{N-m}F(k-1,N). \tag{A.5}$$

Our goal is to show that $F(k,N)\Big/\binom{N}{m+1} \to 0$ as $N \to \infty$.

Consider first the case $k = m$. Since $F(m-1,N) = 0$, (A.5) implies that $F(m,N) = C$ for some constant $C$ independent of $N$. Thus, in particular,

$$H(m,N) - \bar{H}(m) = \frac{C}{\binom{N}{m+1}}, \tag{A.6}$$

and $H(m,N) \to \bar{H}(m)$ as $N \to \infty$. Then when $k = m+1$, (A.5) becomes

$$F(m+1,N+1) = F(m+1,N) + \frac{m}{N-m}C \tag{A.7}$$

and hence

$$F(m+1,N) = F(m+1,m+1) + C\sum_{N'=m+1}^{N-1} \frac{m}{N'-m} \sim \ln N \tag{A.8}$$

for large $N$.

Furthermore, it is possible to show by induction on $k$ that

$$|F(k,N)| \leq C(k)(m\ln N)^{k-m} \tag{A.9}$$

for some $C(k)$ independent of $N$. In other words,

$$\left|H(k,N) - \bar{H}(k)\right| \leq C(k)\frac{(m\ln N)^{k-m}}{\binom{N}{m+1}} \tag{A.10}$$

which approaches 0 as $N \to \infty$.

Appendix B

Behavior of $\langle l|k\rangle$

Here we show that $\langle l|k\rangle^+ \leq \langle l|k\rangle < 2\langle l|k\rangle^+$, where $\langle l|k\rangle^+$ and $\langle l|k\rangle$ are the average longest counter-clockwise (CCW) link length and the average longest link length, respectively, for a node with degree $k$. For simplicity we work exclusively with quantities that are $N$ (time) independent.

Let $X(k,l,l')$ be the fraction of nodes that have degree $k$, longest CCW link length $l$ and longest clockwise (CW) link length $l'$. Then $\langle l|k\rangle$ can be written as

$$\langle l|k\rangle = \frac{\sum_l l\left(\sum_{l'\leq l} X(k,l,l') + \sum_{l'<l} X(k,l',l)\right)}{\sum_{l,l'} X(k,l,l')}. \tag{B.1}$$

The numerator in (B.1) satisfies the following inequality:

$$\begin{aligned}
\sum_l l\left(\sum_{l'\leq l} X(k,l,l') + \sum_{l'<l} X(k,l',l)\right) &= \sum_l l\left(\sum_{l'\leq l} X(k,l,l') + \sum_{l'<l} X(k,l,l')\right) \\
&= 2\sum_l l\sum_{l'\leq l} X(k,l,l') - \sum_l lX(k,l,l) \\
&< 2\sum_l l\sum_{l'\leq l} X(k,l,l') \\
&\leq 2\sum_l l\sum_{l'} X(k,l,l'),
\end{aligned} \tag{B.2}$$

where we have used the fact that $X(k,l,l') = X(k,l',l)$ due to the CCW/CW symmetry in our model. Noting that

$$\frac{\sum_l l\sum_{l'} X(k,l,l')}{\sum_{l,l'} X(k,l,l')} = \langle l|k\rangle^+, \tag{B.3}$$

we have that $\langle l|k\rangle < 2\langle l|k\rangle^+$. In addition, it is clear that the longest link of a node is at least the length of the longest CCW link. Thus, $\langle l|k\rangle^+ \leq \langle l|k\rangle$ and we have our desired result. In particular, this implies that in our model the average longest link for a node, in addition to the average longest CCW (or CW) link, increases exponentially with the node's degree.

Appendix C

Calculation of Model Dimension Spectra

## C.1 The Partition Function Formalism

The partition function formalism [25,26] is an alternative to the fixed sized grid method of Eq. (3.1) for calculating a dimension spectrum of a measure. We demonstrate here how it allows us to relate the various interval and square size distribution functions found in Models 1, 2, 6 and 7 to the dimensions of the point distributions generated by them.

We cover the measure of interest with a disjoint covering $\{S_i\}$, $i = 1, 2...N$, where each element $S_i$ of the covering set has a diameter $\epsilon_i$ less than or equal to $\delta$. (The diameter $\epsilon_i$ is the largest possible distance between two points in $S_i$.) The partition function is defined as

$$\Gamma_q(\tau, \{S_i\}, \delta) = \sum_{i=1}^{N} \mu_i^q / \epsilon_i^\tau, \tag{C.1}$$

where $\mu_i$ is the measure of $S_i$. For a given $\delta$ the covering $\{S_i\}$ is now chosen such that Eq. (C.1) is maximized (for $q > 1$) or minimized (for $q < 1$), which defines

$$\Gamma_q(\tau, \delta) = \begin{cases} \sup_{S_i} \Gamma_q(\tau, \{S_i\}, \delta), & \text{for } q > 1 \\ \inf_{S_i} \Gamma_q(\tau, \{S_i\}, \delta), & \text{for } q < 1 \quad . \end{cases} \tag{C.2}$$

Then, letting $\delta \to 0$, we define

$$\Gamma_q(\tau) = \lim_{\delta \to 0} \Gamma_q(\tau, \delta). \tag{C.3}$$

The quantity $\Gamma_q(\tau)$ experiences a jump from 0 to $+\infty$, as $\tau$ is increased, at a critical value that we denote $\tau(q)$. The dimension $\tilde{D}_q$ of the measure is then defined as

$$\tilde{D}_q = \tau(q)/(q-1). \tag{C.4}$$

In practice, it is difficult to determine whether a particular covering $\{S_i\}$ is optimal in the sense of Eq. (C.2). However, one can often compute the correct dimension by considering specific coverings, in the following sense. Consider a sequence of disjoint coverings $\{S_i^{(m)}\}$ where covering $m$ of the sequence has maximum diameter $\delta^{(m)}$ ($\epsilon_i^{(m)} < \delta^{(m)}$ for every component $i$ of covering $m$), and $\delta^{(m)}$ converges to zero as $m$ approaches infinity. If the limit

$$\hat{\Gamma}_q(\tau) = \lim_{m \to \infty} \Gamma_q(\tau, \{S_i^{(m)}\}, \delta^{(m)}) \tag{C.5}$$

exists, then like $\Gamma_q(\tau)$ it experiences a jump from 0 to $+\infty$, as $\tau$ increases, at a critical value $\tau = \hat{\tau}(q)$. In terms of the sequence $\{S_i^{(m)}\}$ of coverings (not necessarily optimal), we can then compute a dimension spectrum

$$\hat{D}_q = \hat{\tau}(q)/(q-1), \tag{C.6}$$

Since the sequence of coverings used to compute $\hat{D}_q$ may be suboptimal, we have $\hat{\tau}(q) \geq \tau(q)$ for $q > 1$, and $\hat{\tau}(q) \leq \tau(q)$ for $q < 1$. Thus $\hat{D}_q \geq \tilde{D}_q$ in each case. Assuming the limit in Eq. (3.1) that defines $D_q$ to exist, then $D_q$ can be computed by the partition function formalism above using equal size coverings ($\epsilon_i^{(m)} = \delta^{(m)}$ for every $i$ and $m$). Thus $D_q \geq \tilde{D}_q$ as well. While there do exist examples for which $\tilde{D}_q \neq D_q$ for $q < 1$ [1] the two dimensions are typically found to coincide in analytical examples with physical bases. Furthermore, for $q > 1$ the two dimensions always coincide:

**Theorem**: For every probability measure $\mu$ for which the limit in Eq. (3.1) exists, $\tilde{D}_q = D_q$ for all $q > 1$. (If one defines $D_q$ as a liminf, then this equality holds for all $\mu$.)

We prove this theorem in Appendix D.

The theorem demonstrates that, for $q > 1$ at least, the dimension spectrum computed according to the partition function formalism from a particular sequence of coverings is not that sensitive to the type of covering; equal size coverings yield the same spectrum as a sequence of coverings that is optimal in the sense of Eq. (C.2). While it may be possible to compute a different value of $\hat{D}_q$ from a sufficiently suboptimal sequence of coverings, we conjecture that for the coverings we consider below,

$$\hat{D}_q = \tilde{D}_q = D_q. \tag{C.7}$$

[1] As a simple example, we consider the set, $1, 1/2, 1/3, \cdots, n^{-1}, \cdots$, and $q = 0$. For this set it can be shown that Eqs. (C.1)-(C.4), yield $\tilde{D}_0 = 0$, which is intuitively reasonable for a set that is a countable collection of points. In contrast, Eqs. (3.1) and (3.2) yield $D_0 = 1/2$. To see this we first note that, for $q = 0$, (3.1) and (3.2) give the well-known result $D_0 = \lim_{\epsilon \to 0}\{[\ln N(\epsilon)]/[\ln(1/\epsilon)]\}$, where $N(\epsilon)$ is the number of $\epsilon$-intervals needed to cover the set. Next we observe that the distance between $1/n$ and $1/(n+1)$ is approximately $1/n^2$ for large $n$. Thus setting $\epsilon = 1/n^2$, we need $n$ intervals to cover the first $n$ elements of the set, one interval for each such element. To cover the remaining elements we must cover the interval $(1/(n+1), 0)$. This requires $(1/\epsilon)[1/(n+1)] = n^2/(n+1) \simeq n$ intervals. Thus $N(\epsilon) \simeq 2n = 2/\epsilon^{1/2}$, yielding $D_0 = 1/2$.

Furthermore, we conjecture that Eq. (C.7) holds not only for $q > 1$, but for all $q \geq 0$.

For our purposes, the application of Eq. (C.7) for $q \leq 1$ is supported by the following reasoning. In all of the cases below, the formula we obtain for $\hat{D}_q$ is a complex analytic funtion of $q$. If $\hat{D}_q = D_q$ for $q > 1$, it follows that if $D_q$ is continuous at $q = 1$, then $\hat{D}_1 = D_1$. Furthermore, if $D_q$ is complex analytic, then $\hat{D}_q = D_q$ for $q < 1$ as well.

## C.2 Application to the Models

The practical implication of Eq. (C.7) for us is significant. It means that we can choose any sequence of coverings whose maximum diameter converges to zero, and this choice can be made in a manner that *facilitates analytic computations*. The result will be the same as that for an equal cube size covering [Eq. (3.1)] or an optimal covering [Eq. (C.2)].

For Model 1, we choose the intervals between the points as the covering for the distribution. We regard each interval as covering the equivalent of one point since it contains two half points at each of its ends. Hence $\mu_i = (\sum_k N(k,t))^{-1} = 1/t$ for all $i$, and Eq. (C.5) becomes

$$\hat{\Gamma}_q(\tau) = \lim_{t\to\infty} (1/t)^q \sum_{i=1}^{t} \epsilon_i^{-\tau}, \tag{C.8}$$

where the $\delta \to 0$ limit is replaced by $t \to \infty$, since, as $t$ increases the size of the largest interval decreases to zero. Since $\ln \hat{\Gamma}_q(\tau) = -\infty$ for $\tau < \hat{\tau}(q)$ and $\ln \hat{\Gamma}_q(\tau) = \infty$ for $\tau > \hat{\tau}(q)$, we equate $\ln \hat{\Gamma}_q(\tau)$ to zero and obtain $q$ as a function of the transition value $\hat{\tau}$,

$$q(\hat{\tau}) = \lim_{t\to\infty} \frac{\ln \sum_{i=1}^{t} \epsilon_i^{-\hat{\tau}}}{\ln t}. \tag{C.9}$$

In terms of $n(k,t)$, the fraction of the total number of intervals at time $t$ with length $2^{-k}$, we then have [see Eq. (3.9)]

$$q(\hat{\tau}) = 1 + \lim_{t\to\infty} \frac{\ln \langle 2^{k\hat{\tau}} \rangle_t}{\ln t}. \tag{C.10}$$

For Model 2 we use the two types of squares (type I and II) to cover the distribution. A type I square has one point on one of its vertices and hence covers the equivalent of a quarter of a point, while a type II square, with two points on its vertices, covers the equivalent of half a point.

Thus at time $t$, the measure contained in a type I square is $1/4t$, while that contained in a type II square is $1/2t$. Thus Eq. (C.5) becomes

$$\hat{\Gamma}_q(\tau) = \lim_{t\to\infty} \frac{1}{(4t)^q} \left( \sum_{i=1}^{N_1} \epsilon_i^{-\tau} + 2^q \sum_{i=1}^{N_2} \epsilon_i^{-\tau} \right), \tag{C.11}$$

where the first summation is over all type I squares and the second is over all type II squares. Taking the logarithm of this expression and equating it to zero, we obtain

$$q(\hat{\tau}) = \lim_{t\to\infty} \frac{\ln\left(T_1(t) + 2^q T_2(t)\right)}{\ln t}, \tag{C.12}$$

where $T_i(t) = \sum_k 2^{k\hat{\tau}} N_i(k,t)$.

Each square in Model 6 and triangle in Model 7 contains $(\sum_k N(k,t))^{-1} = 1/(3t-2)$ of the total measure. Thus, Eq. (C.5) becomes

$$\hat{\Gamma}_q(\tau) = \lim_{t\to\infty} \frac{1}{(3t-2)^q} \sum_{i=1}^{3t-2} \epsilon_i^{-\tau}. \tag{C.13}$$

We see that this yields an expression for $q(\hat{\tau})$ identical to that for Model 1 [Eq. (C.10)].

Appendix D

Proof of Theorem from Appendix C

As we noted in Appendix C, $\tilde{D}_q \leq D_q$ for all $q \geq 0$, so we now show that also $\tilde{D}_q \geq D_q$ for $q > 1$. Our proof is based on the theorem in [27] that for $q > 1$, the quantity

$$I_q(s) = \int \left( \int \frac{d\mu(y)}{|x-y|^s} \right)^{q-1} d\mu(x) \tag{D.1}$$

is finite for $s < D_q$ and infinite for $s > D_q$. To see that $\tilde{D}_q \geq D_q$, we show that $\tilde{D}_q \geq s$ whenever $I_q(s)$ is finite. Let $\{S_i\}$ be a disjoint covering of $\mu$, as in Appendix C, and recall that $\mu_i$ is the measure of $S_i$ and $\epsilon_i$ is its diameter. Then considering only the contribution to $I_q(s)$ from points $x$ and $y$ that are in the same $S_i$, we have

$$\begin{aligned}
I_q(s) &\geq \sum_{i=1}^{N} \int_{S_i} \left( \int_{S_i} \frac{d\mu(y)}{|x-y|^s} \right)^{q-1} d\mu(x) \\
&\geq \sum_{i=1}^{N} \int_{S_i} \left( \int_{S_i} \frac{d\mu(y)}{\epsilon_i^s} \right)^{q-1} d\mu(x) \\
&= \sum_{i=1}^{N} \mu_i \left( \frac{\mu_i}{\epsilon_i^s} \right)^{q-1} \\
&= \sum_{i=1}^{N} \mu_i^q \Big/ \epsilon_i{}^{(q-1)s} \\
&= \Gamma_q\left((q-1)s, \{S_i\}, \delta\right)
\end{aligned}$$

by (C.1). This implies that $\Gamma_q\left((q-1)s, \delta\right) \leq I_q(s)$ by (C.2), since $q > 1$, and hence $\Gamma_q\left((q-1)s\right) \leq I_q(s)$ by (C.3). Therefore if $I_q(s)$ is finite, then $\Gamma_q\left((q-1)s\right)$ is finite, whence $\tau(q) \geq (q-1)s$, and finally $\tilde{D}_q \geq s$ by (C.4).

BIBLIOGRAPHY


[1] S. Wolfram, *A New Kind of Science* (Wolfram Media, Champaign 2002).

[2] J. M. Epstein and R. L. Axtell, *Growing Artificial Societies: Social Science From the Bottom Up* (MIT Press, Cambridge 1996).

[3] C. Mayhew and R. Simmon, *Astronomy picture of the day: Earth at night*, URL http://antwrp.gsfc.nasa.gov/apod/ap040822.html.

[4] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[5] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, PNAS **97**, 11149 (2000).

[6] M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).

[7] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).

[8] R. Albert and A.-L. Barabási, Science **286**, 509 (1999).

[9] M. Faloutsos, P. Faloutsos, and C. Faloutsos, ACM SIGCOMM **29**, 251 (1999).

[10] R. F. i Cancho, C. Janssen, and R. V. Solé, Phys. Rev. E **64** 046119 (2001).

[11] J. Camacho, R. Guimerà, and L. A. N. Amaral, Phys. Rev. Lett. **88**, 228102 (2002).

[12] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).

[13] H. Jeong, S. P. Mason, A.-L. Barabasi, and Z. N. Oltvai, Nature (London) **411**, 41 (2001).

[14] S. H. Strogatz, Nature (London) **410**, 268 (2001).

[15] S. N. Dorogovtsev and J. F. F. Mendes, Adv. Phys. **51**, 1079 (2002).

[16] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[17] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[18] H. Hentschel and I. Procaccia, Physica D **8**, 435 (1983).

[19] P. Grassberger, Phys. Lett. A **97**, 227 (1983).

[20] S. Yook, H. Jeong, and A.-L. Barabási, PNAS **99**, 13382 (2002).

[21] M. Batty and P. Longley, Fractal Cities (Academic Press, San Diego, 1994).

[22] H. A. Makse, S. Havlin, and H. E. Stanley, Nature **377**, 608 (1995).

[23] S. C. Manrubia, D. H. Zanette, and R. V. Solé, Fractals **7**, 1 (1999).

[24] E. Ott, Chaos in Dynamical Systems (Cambridge University Press, Cambridge, 2002), 2nd ed.

[25] P. Grassberger, Phys. Lett. A **107**, 101 (1985).

[26] T. Halsey, M. Jensen, L. Kadanoff, I. Procaccia, and B. Shraiman, Phys. Rev. A **33**, 1141 (1986).

[27] B. Hunt and V. Y. Kaloshin, Nonlinearity **10**, 1031 (1997).

[28] J. B. Knight, H. M. Jaeger and S. R. Nagel, Phys. Rev. Lett., **70**, 3728 (1993).

[29] H. A. Makse, R. C. Ball, H. E. Stanley and S. Warr, Phys. Rev. E **58**, 3357 (1998).

[30] Y. Oyama, Bull. Inst. Phys. Chem. Res. Jpn. Rep. **18**, 600 (1939); (in Japanese).

[31] K. M. Hill and J. Kakalios, Phys. Rev. E **49**, 3610 (1994).

[32] K. Choo, T. C. A. Molteno and S. Morris, Phys. Rev. Lett., **79**, 2975 (1997).

[33] K. M. Hill, A. Caprihan, and J. Kakalios, Phys. Rev. Lett. **78**, 50 (1997).

[34] K. M. Hill, A. Caprihan, and J. Kakalios, Phys. Rev. E **56**, 4386 (1997).

[35] O. Zik, D. Levine, S. G. Lipson, S. Shtrikman, and J. Stavans, Phys. Rev. Lett. **73**, 644 (1994).

[36] D. Levine, Chaos **9**, 573 (1999).

[37] I. S. Aranson and L. S. Tsimring, Phys. Rev. Lett. **82**, 4643 (1999).

[38] T. Elperin and A. Vikhansky, Phys. Rev. E **60**, 1946 (1999).

[39] T. Shinbrot and F. J. Muzzio, Phys. Today, March issue, 25 (2000).

[40] S. Puri and H. Hayakawa, Physica A **290**, 218 (2001).

[41] T. Yanagita, Phys. Rev. Lett. **82**, 3488 (1999).

[42] S. Das Gupta, D. V. Khakhar and S. K. Bhatia, Chem. Eng. Sci. **46**, 1513 (1991).

[43] M. Newey, J. Ozik, S. M. van der Meer, E. Ott and W. Losert, Europhys. Lett. **66**, 205 (2004).