COMPUTING MULTIFRACTAL SPECTRA VIA SIMPLICIAL MEASURES

by

Jesse James Berwald

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Mathematics

MONTANA STATE UNIVERSITY
Bozeman, Montana

August, 2011

APPROVAL

of a dissertation submitted by

Jesse James Berwald

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to The Graduate School.

Dr. Tomáš Gedeon

Approved for the Department of Department of Mathematical Sciences

Dr. Kenneth Bowers

Approved for The Graduate School

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Jesse James Berwald

August, 2011

DEDICATION

To Sarah.

ACKNOWLEDGEMENTS

It is a pleasure to thank those who have made this dissertation possible. In particular, I extend my gratitude to my advisor, Dr. Tomáš Gedeon. His enthusiasm, excitement, and insight have made this work enjoyable. I also would like to acknowledge my committee, in particular Dr's. Lukas Geyer and Jarek Kwapisz, both of whose time and support over the past few years has been invaluable.

I would like to thank both Scott Mitchell and Afra Zomorodian, who provided great help at Sandia National Laboratory in matters both mathematical and computational.

I am also grateful to the staff in the math department at Montana State University, who keep the machine running smoothly.

I extend my heartfelt gratitude to Sarah Berwald, whose love and support are more appreciated than she knows. I am also deeply grateful to my parents, Susan Murany and Jim and Carole Berwald, who have always supported me, even through some of my stranger decisions. Of course, no acknowledgement would be complete without a special thank you reserved for Ripley Berwald, as well.

Lastly, I would like to extend a special thank you to my friends in Bozeman and beyond: Don, Nikki, Erich, Kyle, T & N, Ian and many others, including all the pups.

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

ABSTRACT

Complex dynamical systems occur on many scales in the natural world, and serve as rich subjects of study. Examples include ecosystems, physiological systems, and financial markets. Simplified versions of these system can be described by dynamical systems. As such, understanding the qualitative behavior of dynamical systems provides an important window into real-world phenomena.

In this manuscript we focus on the qualitative behavior described by the measure concentrated on the attractor of a dynamical system. A common way to study such complicated measures is through their multifractal spectra. We will describe a new method, developed to approximate the Sinai-Bowen-Ruelle measure on an attractor, that is based on the Vietoris-Rips complex. We use it to approximate various measures concentrated on a number of example sets, and demonstrate its efficacy by computing the corresponding multifractal spectra.

CHAPTER 1
INTRODUCTION

Understanding complex systems in the natural world is an increasingly important scientific problem. Mathematics once thought too abstract for use outside of the field has increasingly been harnessed to analyze and solve such real-world problems. Questions about genetics found their answers in combinatorics; number theory encrypts our online transactions; and algebraic topology has been enlisted to help researchers in fields from cardiology to fluid dynamics [51, 54, 71].

A two-pronged approach is used by the scientific community in a steady drive towards a better understanding of the physical world. On one side, measurements taken of real-world phenomena provide raw data with which to gain insight into the system under study. Often, real-world data come in the form of time series in which a single variable in a multi-dimensional system, say temperature, is measured over time. On the other side of the coin are large, complex, mathematical models of real-world systems. From a mathematical perspective these are often dynamical systems. Such models are invariably simplifications of real-world phenomena, yet they provide a controlled environment from which one can gain a deeper understanding of the system.

1.1 Dynamical Systems and Fractals

The study of dynamical systems focuses on the long-term behavior of trajectories in phase space. The field originated with H. Poincaré's study of celestial mechanics in the 1890's. In his investigations of planetary motion and the stability of the solar system he found the first indications that even a small system with three moving

bodies is not structurally stable. * This implies that deterministic nonlinear systems can exhibit very complex long term behavior.

Numerous methods have been developed to understand, in a qualitative manner, the complex behavior of dynamical systems. A qualitative feature of a mathematical object is a value associated to that object that distinguishes it from other objects based on a notion of "type". Topology provides an intuitive example. The fundamental group of the circle has a single generator, a fact which we denote by $\pi(S^1) \cong \mathbb{Z}$ [60]. Contrast this with the fundamental group of the torus, $\mathbb{T}^2$, which has two generators, or equivalently $\pi(\mathbb{T}^2) \cong \mathbb{Z} \times \mathbb{Z}$ [60]. The fundamental groups are the same for any objects homeomorphic to the circle and torus. That is, by reducing the description of an object to a description by its fundamental group we are able to quickly distinguish it from others. Hence, we neglect a precise geometric description of the object in favor of one that gives its principle characteristics. Such characterizations are inherently qualitative and serve to differentiate topological objects.

Similarly, in seeking to understand real-world or modeled dynamical systems we do not look for precise solutions. Rather, we ask more "global", qualitative questions about the topological characterization of the global attractor, and what properties an invariant measure on that attractor might possess. Implicit in these questions is how the parameters of a model affect the dynamics, and therefore the attractor. Another important question is how we bridge the divide between experimental observations of a dynamical system and the theory that describes these systems. Our window into the behavior of the system is through the dimension of the support of a certain invariant probability measure, which is termed a Sinai-Bowen-Ruelle (SBR) measure.

*A system of differential equations (or a map) is structurally stable if small perturbations do not cause a qualitative change in the behavior of the system.

It describes how frequently certain regions of the attractor are visited by the orbit of the dynamical system.

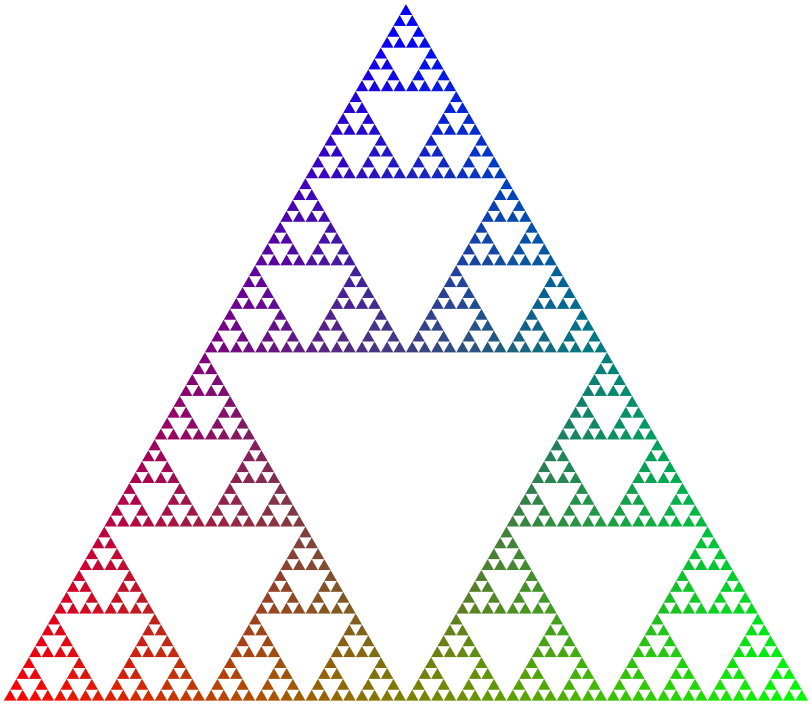


Figure 1.1: The Sierpinski gasket. It has a self-similar structure and a fractal dimension of $\log 3 / \log 2$.

Focusing on the dimension of the attractor leads directly to the notion of a fractal. The word "fractal" is not well-defined, in either the linguistic or the mathematical sense. Therefore, when we say that X is a "fractal" we imply that X is an object

i. whose Hausdorff dimension strictly exceeds its topological dimension;

ii. that has a structure that has detail at arbitrary scales (this may be due to self-similarity properties).

Such sets are not just mathematical curiosities. Highly irregular sets exist in nature. One of the classic examples comes from B. Mandelbrot in his seminal paper,

How Long Is the Coast of Britain? Statistical Self-Similarity and Fractal Dimension [57]. He argued that physical curves such as coastlines are so complex as to be, if not infinitely long, then at least undefinable as functions. He approached the problem through the lens of fractal geometry and characterized such curves by their fractal dimensions.

The intuitive definition of dimension is known more properly as *topological dimension* and can be defined as the number of independent parameters necessary to describe a point in a space. Hence, it is always an integer. Familiar spaces from mathematics and the physical world–circles, tori, $\mathbb{R}^2$–all have integral dimensions. The *Hausdorff* dimension generalizes the notion of dimension to values in $[0, \infty)$. For the regular sets with integral dimensions mentioned above, the topological and Hausdorff dimensions agree. However, this is not always the case, as a classic example of a fractal, the Sierpinski gasket in Figure 1.1, shows. The self-similarity property is clearly visible. The Sierpinski gasket can be described by an iterated function system whose attractor has Lebesgue measure zero and Hausdorff dimension $\log 3 / \log 2$. The implication here is that there is another, more natural measure supported entirely on the attractor. The main thrust of this manuscript is aimed at understanding through experimental means measures supported on fractal sets.

1.2 Contributions

"Topology! The stratosphere of human thought! In the twenty-fourth century it might possibly be of use to someone..."

The First Circle, A. Solzhenitsyn

Data collection from the real-world as well as data generated by mathematical models present a formidable challenge to mathematicians and scientists. Understand-

ing how parameters of a dynamical system can affect changes in that system has important implications for many scientific fields, from climate science to physiology.

In this thesis we are concerned primarily with the invariant SBR measure on an attractor and ways in which such a measure can be approximated. We develop a new method of approximation, which we term the *simplicial measure*, in order to study such measures. The simplicial measure is constructed from the maximal simplices in Vietoris-Rips complexes, which are shown to give an accurate representation of the density of points in a finite data set near the attractor of a dynamical system. Approximation of measures typically proceeds by considering a sequence of covers of decreasing resolutions of a fractal attractor. Each cover contains a collection a minimal "balls" (which merely need to be simply connected regions) intersecting points in the attractor [36]. We do not deviate from this basic notion. Where we diverge is in the formation of the covers. Traditional methods consider covers using grids of side-length r. Thus, a cover is not sensitive to the distribution of points in the attractor. By contrast, a cover of a set using maximal simplices places each element of the collection directly over the regions of highest density.

We apply these techniques to a number of examples by studying their *multifractal spectra*. Using the mathematical examples of the Cantor set and the Hénon attractor we show that the simplicial measure accurately and efficiently approximates the natural SBR measure concentrated on the respective attractors. In particular, our method uses orders of magnitude fewer points than the existing methods to estimate the multifractal spectrum of the SBR measure. This is an important consideration in experimental situations where there is a non-trivial cost associated with acquiring each data point. In addition to these examples, which can be seen as theoretical in nature, we study a computational model of a neural network, as well as heart beat data from healthy and compromised individuals.

The time series produced by the model neural network has previously been shown to exhibit characteristics of a healthy physiological system [17]. The results in Chapter 4 confirm this conclusion. Our results agree very well with previous multifractal spectra computed from the time series of healthy individuals. Additionally, we analyze heart beat time series data for both healthy and compromised patients using the simplicial measure to compute the multifractal spectra. As has been shown previously [49], the range of scales seen in the plots of the multifractal spectra vary greatly depending on the health of the individual. Computing the multifractal spectra using our methods, we show that our results agree extremely well with previous results.

CHAPTER 2
BACKGROUND THEORY

Below we provide an overview of the background necessary to develop the theory and results in the subsequent chapters. We outline the concepts from dynamical systems and dimension theory first. These ideas form the foundation of the results in Chapters 3 and 4. Related to these ideas on the "experimental" side is time series analysis, which we briefly summarize. Finally, we provide a an overview of wavelet theory and point out its application to fractal analysis. Results on fractal analysis derived for wavelets can be applied to the simplicial measure defined in Chapter 3.

2.1 Analysis of Dynamical Systems

Let f be a continuous map from $\mathbb{R}^n$ to itself and suppose that the subset $A \subset \mathbb{R}^n$ is an attractor of f. assume that we do We seek to understand f by studying a particular invariant measure supported on A.

Definition 2.1. An *invariant measure* μ on $\mathbb{R}^n$ is a Borel measure such that

$$\mu(f^{-1}(E)) = \mu(E)$$

holds whenever E is a Borel subset of $\mathbb{R}^n$.

Suppose that A is compact. By the Krylov-Bogolyubov theorem [52] f admits an invariant probability measure supported on A. This result guarantees that we can find at least one invariant probability measure, but does not tell us anything about such a measure. Not all invariant measures admitted by f are relevant. For example, suppose that x_0 is an unstable fixed point of f. The atomic measure δ_{x_0} concentrated

at x_0 is invariant under f. Such a measure tells us little about the attractor or about f [33, 76].

As stated above, we want to understand f through an invariant measure. We require that this measure describes, in a physical sense, the global behavior of f. One way of accomplishing this for a certain class of dynamical systems is by considering a Sinai-Bowen-Ruelle (SBR) measure [70]. Experimentally, an SBR measure is often realized by considering the weak-$*$ limit of time averages of the orbit of an initial point x,

$$\mu = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \delta_{f^i(x)} \tag{2.1}$$

where $x \in \mathbb{R}^n$ and δ_y denotes the unit mass at y. For certain classes of dynamical systems, (2.1) holds for Lebesgue a.e. initial condition x near the attractor of the system.

This definition opens up the possibility to approximate such measure μ by a finite data set. This data takes the form $Y = \{x, f(x), \ldots, f^{N-1}(x)\}$ in the phase space of f. This can be thought of as experimental measurements obtained either from a real-world system or a mathematical model. We assume that N is large enough that the time averages give an accurate picture of the attractor, regardless of any transients. Experimentally, we approximate the measure on A to study the underlying dynamical system f.

2.2 Time Series

Time series occur across a wide variety of fields. They serve as one of the primary windows into the function of a majority of complex dynamical systems around us. Examples include the Dow Jones Industrial Average, average monthly temperatures

in Bozeman, MT, as well as physiological measurements such as inter-beat heart beat intervals [29, 65]. In general, a sequence of measurements taken at equally-spaced* points in time yields a *time series*. The most familiar quantities computed for a given time series process are its mean, variance about the mean, linear regression line, autocorrelation, and various spectral properties [55]. Standard statistical analysis techniques are used to quantify properties of data measurements of real-world systems, as well as to validate models of those systems.

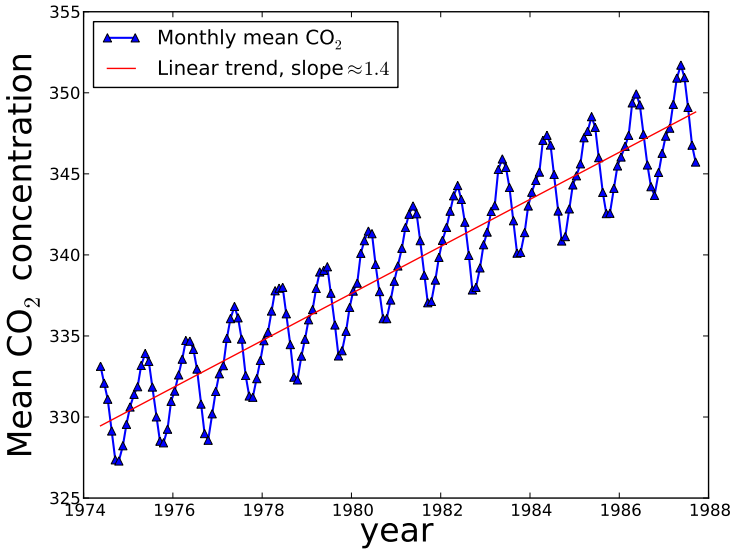


Figure 2.1: A sample time series showing monthly mean CO_2 concentration (mixing ratio in dry air) at the Mauna Loa Observatory in Hawaii, during the period from 1974 to 1987. The best fit linear trend line is plotted, showing the rate of increase to be approximately 1.4. (Source: NOAA, [75])

An example time series is plotted in Figure 2.1. The triangles show the monthly mean CO_2 concentrations measured at Mauna Loa Observatory in Hawaii between 1974 and 1987 are shown Figure 2.1. The rising linear trend is plotted along with

*Time series with measurements obtained at unequally-spaced points in time are allowable, yet a vast majority of the statistical theory surrounding time series assumes equispaced measurements.

the monthly means. Often, real-world time series result from dynamical systems. In such cases, a one-dimensional time series may represent the measurement over time of a single variable (such as temperature), and can afford a window into the larger system [73].

Time series often exhibit fractal behavior over a wide range of scales (see Section 2.3.2). Since time series are typically finite sets, results concerning their fractal properties are necessarily approximations. The analysis of the fractal behavior of time series using wavelets (Section 2.4) was developed extensively during the 1990's by a number of mathematicians and scientists in different fields [2, 10, 38, 56, 61]. Wavelets have a short, but rich history, having been applied to image analysis, data compression, as well as time series analysis. In fractal applications they operate by determining the local Hölder exponents of a function or time series in a given, moving window. By adapting this concept for measures, in Chapter 3 we show how one can analyze the fractal behavior of time series, and more general dynamical systems in higher dimensions, using techniques borrowed from wavelet theory as well as computational topology.

2.3 Measure Theoretic Tools

In what follows we assume familiarity with the basics of measure theory. Appropriate background can be found in [15, 36, 76]. We focus in this section on the aspects specific to the multifractal formalism and dimension theory.

2.3.1 Hausdorff Dimension of a Set

Central to the study of fractals is the concept of the dimension of a set. The definition of dimension most commonly used in dimension theory is the Hausdorff dimension. The exposition herein is based on that of Falconer [36].

Definition 2.2. We list a couple of definitions.

i. Let $X \subset \mathbb{R}^n$ and $\delta > 0$. The diameter of a set U is

$$|U| := \sup\{d(x, y) \mid x, y \in U\},$$

where d is the Euclidean metric on $\mathbb{R}^n$.

ii. A δ-*cover* of X is a countable collection of sets $\{U_i \mid 0 \leq |U_i| \leq \delta\}$ for which

$$X \subset \bigcup_i U_i.$$

For $s \geq 0$, consider the following family of functions on X,

$$\mathcal{H}^s_\delta(X) = \inf_{C \in \mathcal{G}_\delta} \left\{ \sum_i |U|^s \,\middle|\, U \in C \right\}, \tag{2.2}$$

where $\mathcal{G}_\delta$ is the collection of all δ-covers of X. For any sequence $\delta \to 0$, the possible collections available to cover X decrease in number, hence the sequence $\mathcal{H}^s_\delta(X)$ is non-decreasing and the limit exists. The s-dimensional Hausdorff (outer) measure of X is defined as

$$\mathcal{H}^s(X) := \lim_{\delta \to 0} \mathcal{H}^s_\delta(X).$$

Remark 2.3. It can be shown [34, 36] that if X is a Borel subset of $\mathbb{R}^n$ then

$$\mathcal{H}^n(X) = c_n \mathrm{vol}(X), \tag{2.3}$$

where c_n is the inverse of the volume of a unit ball in $\mathbb{R}^n$ and $\mathrm{vol}(\cdot)$ is the volume of X determined using the standard n-dimensional Lebesgue measure. Thus, $\mathcal{H}^s$ agrees with the standard notion of volume, up to a constant factor.

Example 2.4. We show that a continuous, simple curve $g : [a, b] \to \mathbb{R}^n$ has Hausdorff measure L, where L is the arc length of g. Recall that

$$L := \sup \left\{ \sum_{i=1}^{N} |g(t_i) - g(t_{i-1})| \right\}, \tag{2.4}$$

where the supremum is taken over all possible partitions of $[a, b]$, eg., $a = t_0 < t_1 < \cdots < t_N = b$, and N is any natural number. We need the following result whose proof can be found in Falconer [36].

Proposition 2.5. *Let $E \subset \mathbb{R}^n$ and $f : E \to \mathbb{R}^m$ be Hölder continuous such that $\forall x, y \in E$*

$$|f(x) - f(y)| \leq c|x - y|^{\alpha}.$$

Then for each s,

$$\mathcal{H}^{s/\alpha}(f(E)) \leq c^{s/\alpha}\mathcal{H}^s(E).$$

If f is Lipschitz, then $\alpha = 1$ and we get that

$$\mathcal{H}^s(f(E)) \leq c^s\mathcal{H}^s(E). \tag{2.5}$$

We demonstrate first that the Hausdorff measure of $\mathcal{H}^1(g([a, b])) \geq L$. Denote a linear segment in $\mathbb{R}^n$ by $[x, y]$. It can be shown easily that $\mathcal{H}^1([x, y]) = |y - x|$. Let $t_1, t_2 \in \mathbb{R}^n$ such that $a < t_1 < t_2 < b$, and define an orthogonal projection π of g onto the linear segment between $g(t_1)$ and $g(t_2)$ by

$$\pi(g([t_1, t_2]) = [g(t_1), g(t_2)].$$

Note that π is Lipschitz, with $c = 1$. It follows from 2.5 that

$$|g(t_2) - g(t_1)| = \mathcal{H}^1([g(t_1), g(t_2)]) = \mathcal{H}^1(\pi \circ g([t_1, t_2])) \leq \mathcal{H}^1(g([t_1, t_2])).$$

Extending this to any partition of $[a, b]$ gives

$$\begin{aligned}\sum_{i=1}^{N} |g(t_i) - g(t_{i-1})| &= \sum_{i=1}^{N} \mathcal{H}^1([g(t_{i-1}), g(t_i)]) \\ &\leq \sum_{i=1}^{N} \mathcal{H}^1\left(g([t_{i-1}, t_i])\right) \\ &\leq \mathcal{H}^1 g([a, b]).\end{aligned} \tag{2.6}$$

Taking the supremum of the left hand side of 2.6 gives the desired lower bound

$$\mathcal{H}^1\left(g([a, b])\right) \geq L. \tag{2.7}$$

In order to turn the inequality in 2.7 around, consider a parameterization of the arc length $s : [a, b] \to [0, L]$ defined by

$$s(t) = \sup\left\{ \sum_{i=1}^{N} |g(t_i) - g(t_{i-1}) \,\Big|\, a = t_0 < t_1 < \cdots < t_N = t \leq b \right\}.$$

The assumption that g is simple implies that s is strictly increasing, and hence that $s^{-1} : [0, L] \to [a, b]$ exists and is continuous. Now, consider $g \circ s^{-1} : [0, L] \to \mathbb{R}^n$ which parameterizes g by its arc length. This is Lipschitz, so we can apply Proposition 2.5 to get

$$\mathcal{H}^1(g([a, b])) = \mathcal{H}^1(g \circ s^{-1}([0, L])) \leq \mathcal{H}^1([0, L]) = L.$$

We continue with our overview of the Hausdorff measure. Consider two parameters $t > t'$ and choose a δ-cover $\{U_i\}$ of X. Then we get that

$$\sum_i |U|^t \leq \sum_i |U|^{t-t'} |U|^{t'} \leq \delta^{t-t'} \sum_i |U|^{t'}.$$

Taking infima for the above gives $\mathcal{H}^t_\delta(X) \leq \delta^{t-t'} \mathcal{H}^{t'}_\delta(X)$. It follows that, after taking $\delta \to 0$, if $\mathcal{H}^{t'}(X) < \infty$ then $\mathcal{H}^t(X)$ must be zero. Thus, in a graph of $\mathcal{H}^s$ vs. s, at the point $\mathcal{H}^s$ is finite the graph must make a discontinuous jump. The point at which this jump occurs is called the *Hausdorff dimension* of X.

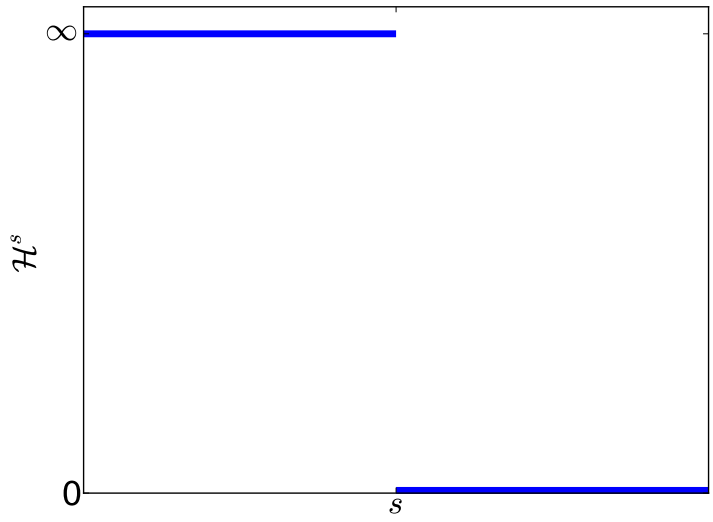


Figure 2.2: A typical graph of $\mathcal{H}^s$.

Definition 2.6. The *Hausdorff dimension* of X is defined as

$$\dim_H(X) := \sup\{s \mid \mathcal{H}^s(X) = \infty\} = \inf\{s \mid \mathcal{H}^s(X) = 0\}.$$

A typical graph of $\mathcal{H}^s$ is shown in Figure 2.2. It is gratifying to confirm that the Hausdorff dimension agrees with our traditional notion of dimension when we consider "normal" sets.

The power of the Hausdorff dimension in analyzing attractors of dynamical systems is that s need not be an integer. The drawback to the Hausdorff dimension is that it can be difficult to compute numerically. We discuss ways around this difficulty in the sections below.

2.3.2 Fractals

As mentioned in the introduction, the word "fractal" does not refer to a single object or even a single category of objects. Therefore, we begin our overview of

fractals with a class of straightforward examples, sets that possesses what is known as *self-similarity* on all scales.

The study of fractal sets hinges on a fundamental axiom of power-law scaling behavior [36]. This quality manifests itself in the following relationship between a "measurement" of X at a given scale and that scale r. If M_r is some measurement of X at scale r, then a *power-law relationship* asserts that there exists some $\alpha \geq 0$ such that

$$M_r(X) \sim r^{\alpha} \iff \alpha = \lim r \to 0 \frac{\log M_r}{\log r} \tag{2.8}$$

On a log-log plot of $\log r$ vs. $\log M_r$ the scaling behavior in (2.8) is linear, from which value the critical exponent α is estimated. The best way to expand on (2.8) is with an example using the Cantor set, E. We show that in this case the scaling exponent α agrees with the Hausdorff dimension $\dim_H(E) = \log 2/\log 3$.

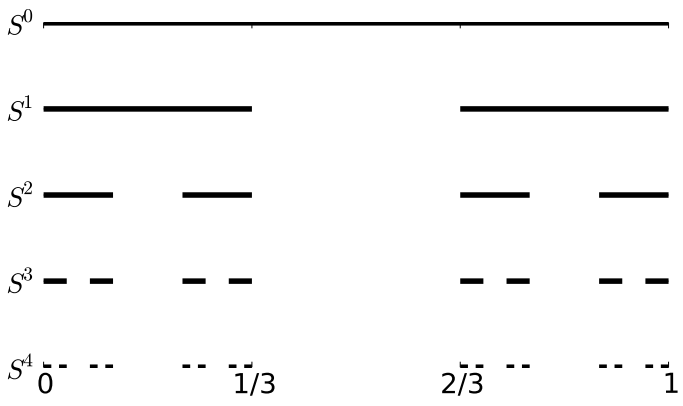


Figure 2.3: The first four iterations in the construction of the ternary Cantor set.

For the set $X \subset \mathbb{R}^n$, define $N_\delta(X)$ to be the number of sets of diameter at most δ necessary to cover X. Thus, N_δ serves as the "measurement" at scale δ.

Example 2.7. We show first that $\dim_H(E) = \log 2/\log 3$. Let E be the ternary Cantor set [36, 52]. The first four stages of the construction have been reproduced in

Figure 2.3. Let E_k be the intervals that make up the k stage. There are 2^k intervals of length $\delta_k = 3^{-k}$ in E_k. The collection E_k is a δ_k-cover of E. Putting this together for the s-dimensional Hausdorff measure, we see that if $s' = \log 2/\log 3$,

$$\mathcal{H}^{s'}_{\delta_k}(E) \leq 2^k \cdot 3^{-ks'} = 1.$$

Thus, taking $k \to \infty$ we get that $\mathcal{H}^{s'}(E) \leq 1$. With a little extra work (see [36]) we get that $\mathcal{H}^{s'}_{\delta}(E) \geq 3^{s'} = 1/2$. Therefore

$$\dim_H(E) = \log 2/\log 3. \tag{2.9}$$

Now, a path to (2.9) that is more amenable to computation is the following. Consider Equation (2.8) in the current context. For $r = \delta_k = 3^{-k}$, there are

$$N_{\delta_k}(E) = 2^k = \delta_k^{\log 2/\log 3},$$

intervals necessary to cover E. Taking logs, we get the scaling behavior shown in the log-log plot of Figure 2.4, with slope $\alpha = \log 2/\log 3$. Hence, for the Cantor set, $\alpha = \dim_H(E) = \log 2/\log 3$.

What we just computed is the *box-counting dimension* of the Cantor set. The box-counting dimension is defined for the set $X \in \mathbb{R}^n$ by

$$\dim_B X = \lim_{\delta \to 0} \frac{\log N_\delta(X)}{-\log \delta}, \tag{2.10}$$

when this limit exists. Experimentally, boxes or grids are typically utilized. Given the power law assumption (2.8), the scaling behavior can be approximated by computing the slope of the linear regression line to the graph of $-\log \delta$ vs. $\log N_\delta$, as in Figure 2.4. For the Cantor set, and many other "regular" fractals such as the Koch curve or Sierpinski gasket, the box counting dimension is equal to the Hausdorff dimension

as shown above. In general, we have the following inequality relating Hausdorff and box-counting dimensions [36],

$$\dim_H X \leq \dim_B X. \tag{2.11}$$

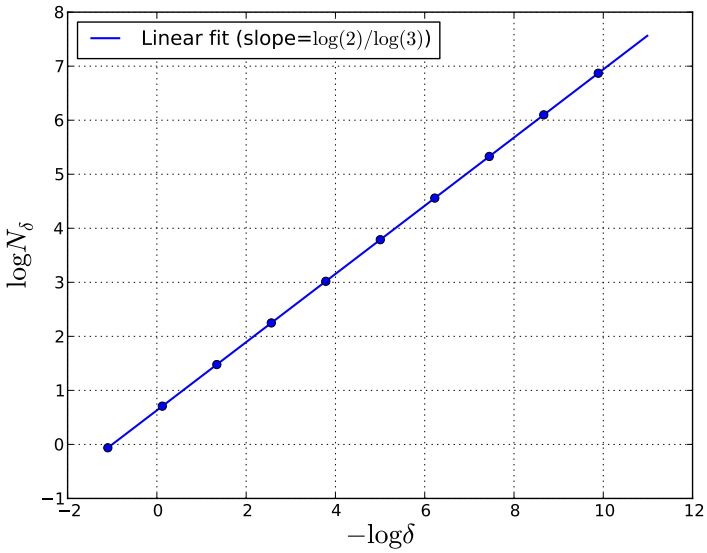


Figure 2.4: The box counting dimension of the Cantor set can be estimated empirically using the slope of a line fit to the set of points.

As we will see in Subsection 2.3.4, the partition function provides another way to access the dimension of a set. As with the box counting dimension, it utilizes observable measurements and so is better-suited to experimental data.

2.3.3 Self-Similar Systems

The dimensions of a few classes of fractals are both analytically and experimentally computable. As we will see, analytically tractable fractals possess a common self-similar structure. These systems are often called *multiplicative processes*. Because we need to compare our algorithms in later chapters with known results, we review a couple of these fractals below.

2.3.3.1 Iterated Function Systems: One of the simplest and most familiar ways to construct a fractal is to define an *iterated function system* (IFS). The Cantor set is the canonical example of an IFS.

An IFS is a set of similarity transformations $\{S_1, \ldots, S_m\}$ where for each $i \in \{1, \ldots, m\}$,

$$d(S_i(x), S_i(y)) = c_i d(x, y), \tag{2.12}$$

where $0 < c_i < 1$ and $x, y \in \mathbb{R}^n$. Define S to be the union of images of the contractions acting on a set so that

$$S(E) = \bigcup_{i=1}^{m} S_i(E),$$

where E is compact in $\mathbb{R}^n$. Then we can define $F \subset E$ to be the unique, nonempty, compact *attractor* of the IFS [36, 52]:

$$F = \bigcap_{i=0}^{\infty} S^i(E), \tag{2.13}$$

Example 2.8. The Cantor set can be constructed in this manner. The specific contractions necessary to arrive at the ternary Cantor set are

$$S_1 = \frac{x}{3} \text{ and } S_2 = \frac{x}{3} + \frac{2}{3}, \tag{2.14}$$

where the contraction ratio are $c_1 = c_2 = 1/3$.

2.3.3.2 Mass Distributions: A mass distribution provides a specific realization of a measure that is simple to construct both theoretically and experimentally. We provide an overview of the key concepts below. In Chapter 3 we utilize mass distributions to arrive at theoretical results concerning the simplicial measure. In Chapter 4 we compute experimental results using approximates of mass distributions.

The general idea of a mass distribution (MD) is simple and involves repeatedly subdividing a mass amongst parts of a (bounded) Borel set E [36]. They are related the IFS's above. For instance, one can imagine an IFS as providing the "shell" into which we place the subdivided mass.

Without loss of generality, take a bounded set $E_0 \subset \mathbb{R}^m$ and a Borel measure μ such that $\mu(E_0) = 1$. Let $\mathcal{E}_0 = \{E_0\}$ be the collection containing the single set E_0. For $k \in \mathbb{N}$, define $\mathcal{E}_k$ to be a collection of disjoint Borel sets $E \subset E_0$ such that for every $E \in \mathcal{E}_k$,

i. E is contained in one of the set in $\mathcal{E}_{k-1}$;

ii. E contains a finite number of sets from $\mathcal{E}_{k+1}$.

We assume that the diameters of the sets in $\mathcal{E}_k$ tend to zero as $k \to \infty$. Additionally, the mass is split amongst the sets $E \in \mathcal{E}_k$ in the following manner. Proceeding from the zeroth to the first level, let $\mathcal{E}_1 = \{E_1, \ldots, E_{n_1}\}$. Then,

$$\sum_{i=1}^{n_1} \mu(E_i) = \mu(E_0).$$

In general, if $\{E_1, \ldots, E_{n_{k+1}}\} \subset \mathcal{E}_{k+1}$ which are contained in the set $E \in \mathcal{E}_k$, then

$$\sum_{i=1}^{n_{k+1}} \mu(E_i) = \mu(E).$$

Let $B_k = \bigcup_{i=1}^{n_k} E_i$,where $E_i \in \mathcal{E}_k$, and define $\mu(U) = 0$ whenever $U \cap B_k = \emptyset$. Then there exists a unique measure μ with these properties that is supported on $\bigcap_{i=0}^{\infty} B_i$ [36]. Figure 2.5 shows this process for a subdivision consisting of two subsets at each stage. The top ellipse is $E_0 = \mathcal{E}_0$, the middle contains $\mathcal{E}_1 = \{E_1, E_2\}$, with the next level $\mathcal{E}_2$ consisting of $E_{11}, E_{12} \subset E_1$ and $E_{21}, E_{22} \subset E_2$. For instance, $\mu(E_1) + \mu(E_2) = p_1 + p_2 = \mu(E_0)$.

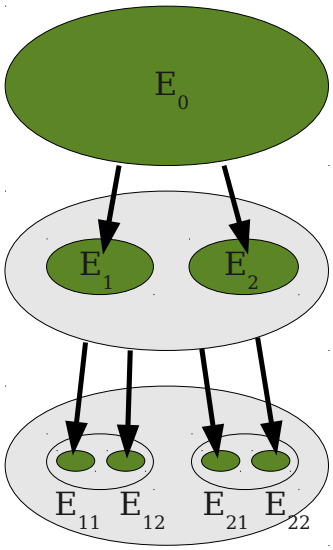


Figure 2.5: The mass (shaded/green) is conserved as it is recursively distributed to a nested collection of sets $\mathcal{E}_k$.

The following example constructs the Cantor set in the context of a mass distribution and will serve as a canonical construction in the theory and applications that follow in Chapters 3 and 4.

Example 2.9. The ternary Cantor set is a special case of the above construction on the unit interval. At the kth stage of the construction the contractions splits the preceding interval is split into two subintervals $^1/_3$ the length of the original. Canonically, the mass is split evenly between the two subintervals.

We may also construct the Cantor set by splitting a mass unevenly at each stage, while leaving the contractions of the intervals unchanged. This results in a measure on the Cantor set whose mass distribution is nonuniform. An example with *mass ratios* $p_0 = 0.2$ and $p_1 = 0.8$ is shown in the histogram in Figure 2.6. We define mass ratios in Equation 2.16.

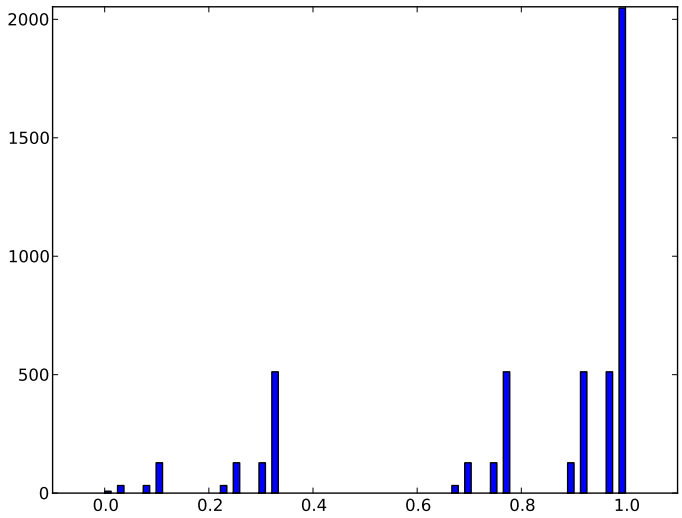


Figure 2.6: The Cantor set, approximated to the 4th level (S^4). The mass ratios are $p_1 = 0.2$ and $p_2 = 0.8$. The figure shows a histogram of 5000 points distributed across the 2^4 intervals according the mass ratios.

2.3.3.3 Binomial Measures: In this subsection we focus on a special case of mass distributions known as binomial measures. Many fruitful examples in dimension theory arise from these objects. The study of binomial measures, sometimes referred to as binomial cascades, has ranged from theoretical examples found in [13, 36, 58, 67], to models of fully developed turbulence [59] and human heartbeat analysis [50]. The binomial measure is fundamental to the result in Chapter 3 concerning the existence of a wavelet transform whose coefficients are approximated by the simplicial measure. In particular, the scaling of the intervals aligns with the traditional dyadic scaling of the analyzing wavelets.

A binomial measure is a special case of a general mass distribution by specifically dividing the mass into two portions at each step. In order to construct a mass distribution we recursively distribute mass at each stage of the construction to a

nested sequence of decreasing intervals. The first few steps of constructing a binomial measure are shown in Figure 2.5.

We construct a deterministic binomial measure on $[0,1]$ from nested dyadic intervals,

$$I_{k_n}^n = [k_n 2^{-n}, (k_n+1)2^{-n}), \tag{2.15}$$

where $n \in \mathbb{N}$ and $k_n \in \{0, 1, \ldots, 2^n - 1\}$. The following assumptions on the measures are stronger than for general mass distributions:

i. $p_0 + p_1 = 1$

ii. The mass ratios are identical at all levels:

$$p_k = \begin{cases} p_0, & \text{if } k \text{ odd} \\ p_1, & \text{if } k \text{ even,} \end{cases} \tag{2.16}$$

for all n and k.

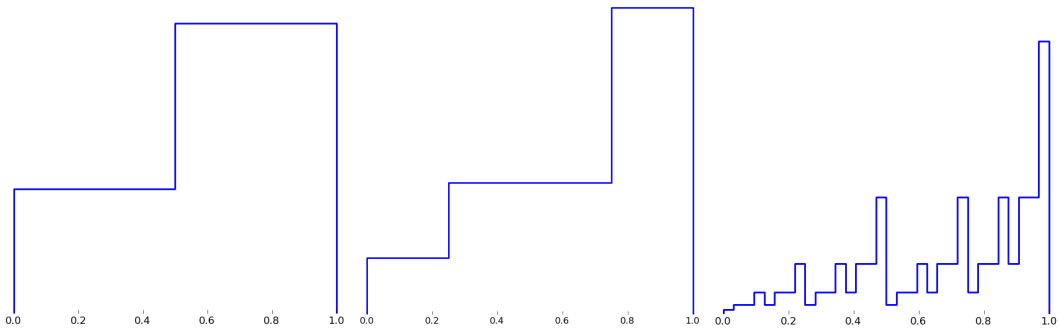
The dyadic tree in (2.18) shows the first few levels of the partitioning of $[0,1]$ from a schematic point of view. In general, the node $I_{k_n}^n$ is connected to the two nodes $\{I_{2k_n}^{n+1}, I_{2k_n+1}^{n+1}\}$ which are located one level "lower". We track the distribution of mass in a parallel manner. As in the general case of a mass distribution, we define the measure μ to be the mass supported on the intervals $I_{k_n}^n$. Thus, if $I_{k_n}^n$ is an interval at the n stage of the construction,

$$\mu(I_{k_n}^n) = p_{k_n} \cdots p_{k_1}$$

and

$$\mu(I_{k_n}^n) = \mu(I_{2k_n}^{n+1}) + \mu(I_{2k_n+1}^{n+1}) \tag{2.17}$$

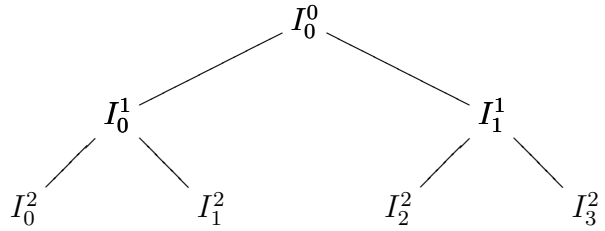
In Figure 2.7 we have plotted the first, second, and fifth levels in the construction of a binomial measure with fixed (deterministic) mass ratios of $p_0 = 0.3$ and $p_1 = 0.7$.



(a) Level 1 (b) Level 2 (c) Level 5

Figure 2.7: The levels 1,2, and 5 of the construction of the binomial measure on $[0, 1]$. Constant mass ratios were used. (The vertical axis has been exaggerated in (b) and (c) to highlight the dyadic structure.)

The height of the horizontal line over each interval represents the relative measure of that interval.



(2.18)

Remark 2.10. The process used to construct a binomial measure is identical to that mentioned in Example 2.9 to construct an uneven mass distribution on a Cantor set, except that the contraction maps are $S_0 = \frac{x}{2}$ and $S_1 = \frac{x+1}{2}$. Therefore, we can transform each interval $I_{k_n}^n$ into a corresponding interval in the ternary Cantor set by composing the respective contractions with the transformations

$$T_0 = \frac{2x}{3} \text{ and } T_1 = \frac{2}{3}x + \frac{1}{3}. \tag{2.19}$$

By applying 2.19 to the left and right intervals at each stage of the construction of a binomial measure we can transform the set that supports the measure.

2.3.4 Multifractal Measures

A *multifractal measure* μ on a set X is a measure whose mass distribution is spread over X is a highly nonuniform manner. In this case, we can decompose X into level sets, parametrized by a scaling exponent parameter α. In particular, let μ be a positive, finite measure on $\mathbb{R}^n$ and define the *local dimension* of μ at $x \in \mathbb{R}^n$,

$$\dim_{\text{loc}} \mu(x) = \lim_{\varepsilon \to 0} \frac{\log \mu(B(x, \varepsilon))}{\log \varepsilon}, \tag{2.20}$$

when it exists. This is sometimes referred to as the *Hölder exponent* at x. (We will return to Hölder exponents in the context of wavelet in Section 2.4.) Using Equation (2.20) define the decomposition of X into the level sets

$$X_\alpha = \{x \mid \dim_{\text{loc}} \mu(x) = \alpha\},$$

where $\alpha \in [\alpha_{\min}, \alpha_{\max}] \subset \mathbb{R}$. When for every $x \in X$, $\dim_{\text{loc}} \mu(x) = \alpha$ then μ is called *monofractal*. An example is an IFS with equal mass distributions and equal contractions. We will sometimes say that X is a "multifractal set", by which we mean that the experimental SBR measure determined by analyzing the ergodic averages on X is multifractal. The multifractal analysis of μ proceeds by computing the Hausdorff dimension X_α for each α:

$$f(\alpha) := \dim_H X_\alpha. \tag{2.21}$$

The dimensions $f(\alpha)$ over $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ comprises the *multifractal spectrum*, sometimes called the *singularity spectrum*, of the measure on X. In Figure 2.8 we plot the multifractal spectra for a Cantor set with a measure constructed from mass ratios $p_0 = 0.2$ and $p_1 = 0.8$, as well as the canonical Cantor set with monofractal mass distribution given by $p_0 = p_1 = 0.5$. The familiar dimension of the

monofractal measure on Cantor set is recovered in the blue circle located at the point $(\log 2/\log 3, \log 2/\log 3)$. The multifractal Cantor set, with a multifractal measure constructed from the mass ratios $p_0 = 0.2$ and $p_1 = 0.8$, shows the wider range of scaling expected. When $q = 0$ the apex of the curve reaches $f(\alpha(0)) = \log 2/\log 3$, where the scaling exponent is $\alpha(0) \approx 0.83$. the

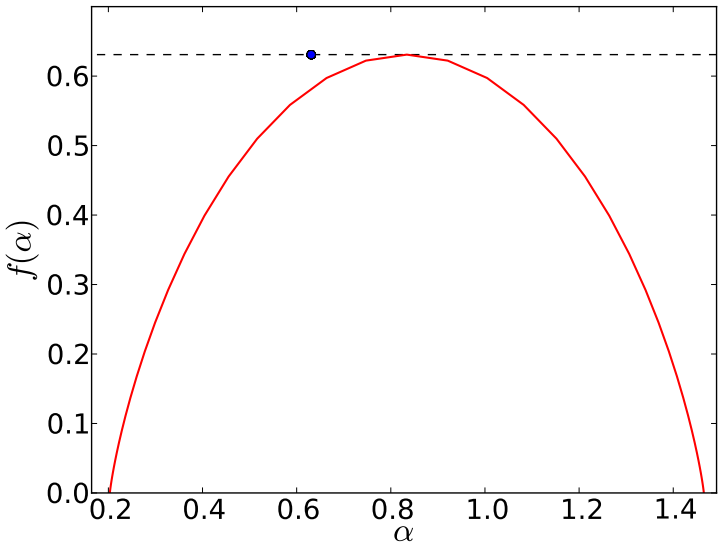


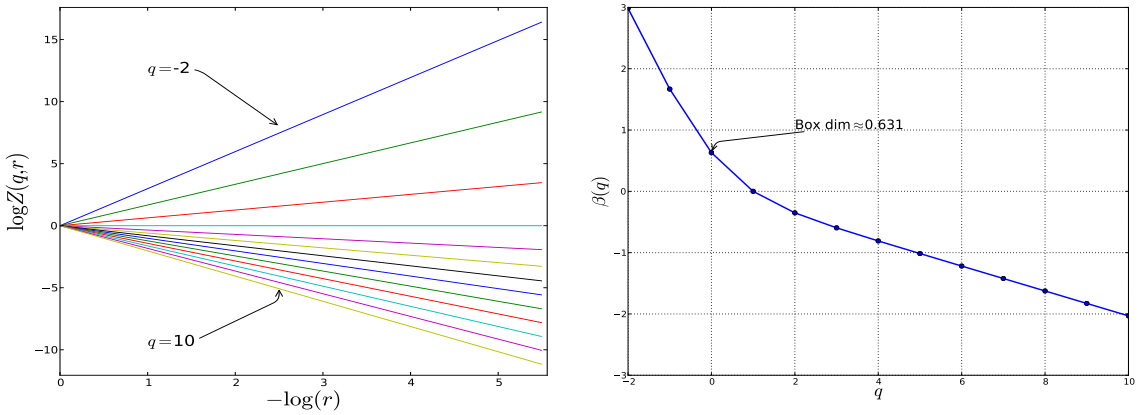
Figure 2.8: The multifractal spectrum for the Cantor set constructed with mass ratios $p_0 = 0.2$ and $p_1 = 0.8$. The dashed line is at a value of $\log 2/\log 3$.

2.3.5 Computing the Multifractal Spectrum

For the purpose of analyzing data from the invariant measure on dynamical system, Equations (2.20) and (2.21) provide limited traction. We summarize two principle methods developed to tackle the problem of computing the multifractal spectrum of multifractal sets. They are closely related. We utilize both methods in Chapter 4 and so review them both below.

As far as we are aware, the only rigorous results connecting theoretical and numerical predictions of multifractal spectra concern multiplicative processes such as

binomial measures and iterated function systems [36, 67]. For the case of wavelets, summarized in Section 2.4, rigorous results seem to be limited to the one-dimensional case for recursively-defined structures [62]. Nevertheless, it is standard practice to apply the techniques below to more complex, higher-dimensional sets such as the Henon attractor and various real-world data [3, 7, 42, 48, 69].



(a) Analytically-determined Partition Functions (b) Analytically-determined $\beta(q)$ Curve

Figure 2.9: (a) Partition functions $Z(q, \varepsilon)$ for $q \in \{-2, -1, \ldots, 10\}$. (b) The $\beta(q)$ curve determined by computing the linear regression slope to each $Z(q, \varepsilon)$ line.

2.3.5.1 The $\beta(q)$ curve: Let X be a compact set in $\mathbb{R}$ that arises through a multiplicative process such as an IFS or mass distribution. We begin by reviewing the moment sum method [36]. Let $\varepsilon > 0$ and fix $q \in \mathbb{R}$. Suppose that $\mathcal{B}_\varepsilon$ is a disjoint cover of X by balls of radius ε. Experimentally, this is often taken to be a set of grid elements of side-length ε. Define the *partition function*

$$Z(q, \varepsilon) = \sum_{B \in \mathcal{B}_\varepsilon} \mu^q(B). \tag{2.22}$$

The parameter q, related to temperature in the thermodynamic version of the partition function, acts to highlight regions of X of intense measure. It is assumed

that $Z(q,\varepsilon)$ also exhibits power-law behavior with respect to ε. Therefore, we write

$$Z(q,\varepsilon) \sim \varepsilon^{-\beta(q)},$$

and examine the scaling properties in the familiar way, by taking logarithms of both sides. For each q, assuming the limit exists, we consider,

$$\beta(q) = \lim_{\varepsilon \to 0} \frac{\log Z(q,\varepsilon)}{-\log \varepsilon}. \tag{2.23}$$

The approximation of $\beta(q)$ is determined by computing the log-log linear regression to the $Z(q,\varepsilon)$ curve in Figure 2.9(a). We find that monofractal measures on a set yield $\beta(q)$ curves that are linear, while multifractal measures yield nonlinear $\beta(q)$ curve. The $\beta(q)$ for a multifractal measure on the Cantor set E is plotted in Figure 2.9(b). Mass ratios of $p_0 = 0.2 = 1-p_1$ were used to construct E. Since the mass distribution can be determined analytically in this case, the values of $Z(q,\varepsilon)$ curves at each ε are exact. Note that when $q = 0$ the computation of the partition function reduces to the box counting dimension of the set. As the annotation in Figure 2.9(b) points out, $\beta(0) = \log(2)/\log(3) \approx 0.631$, the fractal dimension of the monofractal measure on the Cantor set.

Falconer [36] shows that the Legendre transform of $\beta(q)$ is equivalent to the multifractal spectrum $f(\alpha)$. If one desires to compute the $f(\alpha)$ singularity spectrum from experimental data, then explicit smoothing of the β curve is necessary. This requires fitting an nth-degree polynomial to the β curve prior to computing the Legendre transform. The method we describe below provides a way around this annoyance.

2.3.5.2 Implicit determination of $f(\alpha)$: There is an alternative method for computing the $f(\alpha)$ spectrum, first reported by Chhabra [23]. The method computes the curve $f(\alpha)$ from the two curves, $\alpha(q)$ and $f(q)$. Let ν be a measure on $\mathbb{R}$. Suppose

that we are interested in a compact set $X \subset \mathbb{R}$ constructed through a multiplicative process.

The methods detailed here rely on a theorem of Billingsley [18] that, to our knowledge, holds only for sets in $\mathbb{R}$. Billingsley showed that for a multiplicative process the Hausdorff dimension of the measure-theoretic support of ν, say $\mathcal{A}$, is directly related to the entropy of the measure. Let $\mathcal{B}_\varepsilon$ be a finite, disjoint ε-cover of $\mathcal{A}$, then

$$\dim_H \mathcal{A} = \lim_{\varepsilon \to 0} \frac{1}{\log \varepsilon} \sum_{B \in \mathcal{B}_\varepsilon} \nu(B) \log \nu(B). \tag{2.24}$$

We will utilize this result below. For now, let $q \in \mathbb{R}$ and fix a resolution ε. Then we consider the Gibbs measure

$$\begin{aligned} \mu_i(q, \varepsilon) &= \frac{\nu^q(B_i)}{\sum_i^n \nu^q(B_i)} \\ &= \frac{\nu^q(B_i)}{Z(q, \varepsilon)}, \end{aligned} \tag{2.25}$$

where we have enumerated the $B \in \mathcal{B}_\varepsilon$ from $i = 1, \ldots, N$, and $Z(q, \varepsilon)$ is the partition function. This is essentially the Gibbs measure at different "temperatures", where q is analogous to the Boltzmann temperature $1/kT$, and k is the Boltzmann constant [2, 23]. The thermodynamic interpretation provides a nice way to connect (2.25) to the singularity strength α. Suppose that given a resolution ε, the ith "state" of the dynamical system is given by the ith ball in $\mathcal{B}_\varepsilon$. Define the random variable

$$\mathbb{X}_\varepsilon : \Omega \to \mathbb{R},$$

by

$$\mathbb{X}(i) = \alpha_i(\varepsilon) - \frac{\log \nu(B_i)}{\log 1/\varepsilon},$$

where $\Omega_\varepsilon = \{1, 2, \ldots\}$ and $B_i \in \mathcal{B}_\varepsilon$. For $q \in \mathbb{R}$, the probability of observing the ith state is given by the normalized Gibbs measure $\mu_i(q, \varepsilon)$ [23, 67]. Furthermore, if we continue the thermodynamic analogy and suppose that the random variable takes on a value that is identified with the energy per unit volume, then we can define the energy of the ith state as [2, 23].

$$E_i = \alpha_i(\varepsilon). \tag{2.26}$$

Equation (2.26) can be interpreted as the expected energy of the ith configuration at Boltzmann temperature q. See [19, 35] for a summary of the connection between thermodynamics, dynamics, and multifractal analysis.

For each q we can consider the average singularity strength, or average energy per unit volume, $\langle\alpha\rangle$,

$$\langle\alpha\rangle(q) = \lim_{\varepsilon\to 0} \sum_i \mu_i(q, \varepsilon)\alpha_i(\varepsilon) = \lim_{\varepsilon\to 0} \sum_i \mu_i(q, \varepsilon) \frac{\log \nu(B_i)}{\log \varepsilon}. \tag{2.27}$$

Remark 2.11. Notice that 2.27 is related to $\beta(q)$ through its derivative as follows. Assume that the limit exists; the smoothness of $\beta(q)$ is proved in Pesin [67]:

$$\begin{aligned} \frac{\partial\beta}{\partial q} &= \lim_{\varepsilon\to 0} \frac{1}{\sum_i \nu(B_i)^q} \sum_i \nu(B_i)^q \frac{\log \nu(B_i)}{\log \varepsilon} \\ &= \lim_{\varepsilon\to 0} \sum_i \mu_i(q, \varepsilon) \frac{\log \nu(B_i)}{\log \varepsilon}. \end{aligned}$$

Using (2.24) allows us to determine the Hausdorff dimension of the support of the measure $\mu(q)$. By (2.24) and (2.25) this is equivalent [23] to the following

$$f(q) = \dim_H \mathcal{A}_q = \lim_{\varepsilon\to 0} \frac{1}{\log \varepsilon} \sum_{i=1}^{\#\mathcal{B}_\varepsilon} \mu_i(q, \varepsilon) \log \mu_i(q, \varepsilon), \tag{2.28}$$

where $\mathcal{A}_q$ is the measure-theoretic support of $\mu(q)$. Thus, by combining Equations (2.27) and (2.28) we are able to obtain $f(\alpha)$ as an implicit function of q. The spectra in Figure 2.8 were determined in this way.

2.4 Wavelets

The purpose of this section is two-fold. First, we aim to define wavelets in general. Since the use of wavelets to analyze fractal signals and time series is likely the most successful technique to date it is important to understand what wavelets are and the information they extract from functions. Second, wavelet analysis can be applied to measures. This forms a bridge between the simplicial measures described in Chapter 3 and previous results for wavelets concerning multifractal analysis [62].

We start by giving a general overview of wavelets. For a more detailed development see Daubechies [31].

Definition 2.12. A *wavelet* is a function, ψ, with the following properties:

i. $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

ii. $\int_{\mathbb{R}} \psi(x)dx = 0$

iii. $\int_{\mathbb{R}} |\psi(x)|^2 dx = 1.$

The term wavelet (*ondelette*, or small wave in French) is an apt phrase since ψ is both localized in space and of finite energy. An assumption of compact support for ψ would be desirable, it is more convenient to weaken this to property (i) [31]. Wavelets are typically built from a prototypical function, or *mother wavelet*. Let ψ be a mother wavelet for which properties (i)-(iii) hold. The *daughter wavelets* of ψ,

for all $a \in \mathbb{R}^+$ and $b \in \mathbb{R}$, are

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right). \tag{2.29}$$

A complete, orthonormal basis for $L^2(\mathbb{R})$ can be constructed through the process of translation and dilation of the *mother wavelet* [31]. Similarly to the Fourier transform, the *wavelet transform* of a function $f \in L^2(\mathbb{R})$ is defined as,

$$W_\psi f(a,b) = \int_{\mathbb{R}} f(x)\bar{\psi}_{a,b}(x)dx,$$

where $\bar{\psi}$ signifies complex conjugation.

Since the wavelets are localized in space, the decomposition of a function $f \in L^2(\mathbb{R})$ takes place over the scale-space half-plane $\mathbb{R}^+ \times \mathbb{R}$,

$$f(x) = \frac{1}{C_\psi}\int_0^\infty \int_{-\infty}^\infty W_\psi f(a,b)\psi_{a,b}(x)db\,\frac{da}{a^2}, \tag{2.30}$$

where C_ψ is a normalization constant that depends on the particular wavelet ψ.

2.4.1 Multiresolution Analysis

We turn now to the decomposition of functions using discrete wavelets. Discrete wavelets are analogous to the continuous wavelets defined in (2.29), except that translation and dilation are performed using discrete (usually dyadic) factors.

A *multiresolution analysis* serves as a method for constructing discrete wavelets using a convenient, recursive structure. These form the analyzing functions in a discrete analogue of the wavelet transform, and by extension we obtain a discrete version of the decomposition of f in 2.30.

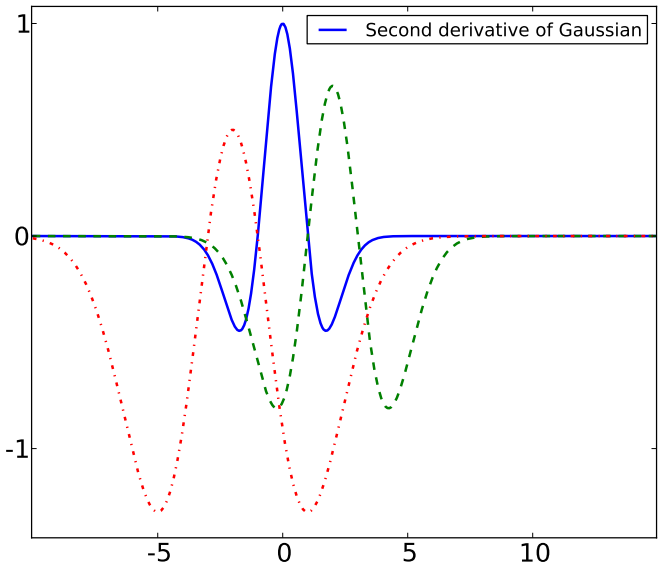


Figure 2.10: The second derivative of a Gaussian used for the mother wavelet ψ (solid, blue line). The dashed functions show two versions of scaled and shifted daughter wavelets.

Formally, multiresolution approximation consists of a sequence $\{V_j\}_{j\in\mathbb{Z}}$ of closed subspaces of $L^2(\mathbb{R})$ such that

$$\{0\} \cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots \tag{2.31a}$$

$$\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$$

$$\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$$

$$\forall (j,k) \in \mathbb{Z}^2,\ \phi(x) \in V_j \iff \phi(x - 2^j k) \in V_j \text{ (translation)}$$

$$\forall j \in \mathbb{Z},\ \phi(x) \in V_j \iff \phi\left(\frac{x}{2}\right) \in V_{j+1} \text{ (scaling)}.$$

The purpose of the approximation spaces V_j is to represent a data set (in general, a collection of vectors in $\mathbb{R}^n$) as a one-parameter family of "smoothed" data. Typically, in applications such as image processing, local maxima or minima that occur on a

scale less than $\sqrt{2^{-j}}$ are averaged out when analyzed with a wavelet from V_j. This has the effect of not only smoothing the image but also downsampling it. An example of such smoothing is shown in Figure 2.11. The original image in Figure 2.11(a) has been approximated in Figure 2.11(b) by a Haar scaling function $\phi_{1,k} \in V_1$. Notice that the smallest features below the resolution of the wavelet (strands of hair, small tree branches, the ridge line) have been "averaged out" into the surrounding, lighter background.

(a) Original Image (b) Approximation Image

Figure 2.11: (a) Original greyscale image. (b) Approximation using the Haar scaling functions $\phi \in V_{-1}$ with scaling 2^{-1} smooths the edges and produces an image of half the size of the original (240×180 pixels versus 480×360 pixels).

In what follows we will focus on the Haar multiresolution analysis, since we make use specifically of the Haar scaling and wavelet functions in Chapter 3. The Haar wavelet is the oldest orthogonal wavelet, created by Alfréd Haar in 1910 to serve as a countable system of orthogonal basis functions for the $L^2(\mathbb{R})$ [44]. In addition to the traditional Haar system, whose elements are synonymous with the Haar wavelets defined below, there exists the corresponding system of Haar scaling functions. Together, the scaling functions and wavelets constitute a simple, yet powerful multiresolution analysis [31].

In the case of a Haar multiresolution analysis we begin with the 0th level scaling functions. (These are also referred to as "father wavelets" by Daubechies [31].)

$$V_0 = \{g \in L^2(\mathbb{R}) \mid g|_{[k,k+1)} = \text{constant}, \, k \in \mathbb{Z}\},$$

The orthonormal basis of V_0 is composed of the scaling functions

$$\{\phi_{0,k} \mid k \in \mathbb{Z}\},$$

where the $\phi_{0,k}$ are shifted indicator functions,

$$\phi_{0,k}(x) = \phi(x-k) = \begin{cases} 1 & \text{if } k \le x < k+1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.32}$$

For all $j \in \mathbb{Z}$, the corresponding orthonormal approximation spaces are

$$V_j = \{g \in L^2(\mathbb{R}) \mid g|_{[k2^j,(k+1)2^j)} = \text{constant}, k \in \mathbb{Z}\},$$

and the orthonormal basis functions for V_j are

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k), \tag{2.33}$$

where $\phi := \phi_{0,0}$ and $k \in \mathbb{Z}$.

Now we describe how the Haar mother wavelet, $\psi = \psi_{0,0}$, is derived from the scaling functions. (Further details can be found in Daubechies [31].) The ladder of subspaces in (2.31a) allows the definition of the wavelet space W_j, which is a subspace of V_{j+1} that satisfies

$$V_{j+1} = V_j \oplus W_j. \tag{2.34}$$

For $f \in L^2(\mathbb{R})$ define

$$D_{j,k} = \int f \phi_{j,k}.$$

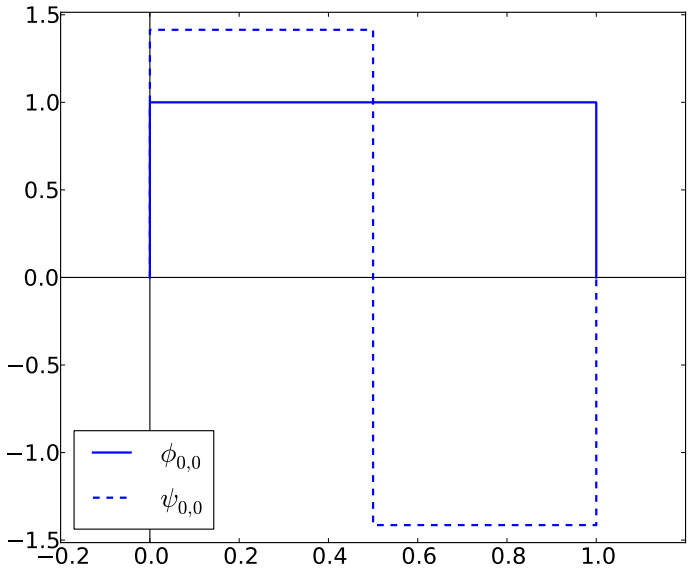


Figure 2.12: The basic Haar scaling function $\phi_{0,0}$ (solid). The Haar mother wavelet $\psi_{0,0}$ (dashed) is the linear combination of the two Haar basis functions for the approximation space V_{-1}.

From (2.31a), it is clear that for any $\tilde{\phi} \in V_0$,

$$\tilde{\phi} = \sum_k D_{1,k}\phi_{1,k}, \tag{2.35}$$

where $\sum_k |D_{1,k}|^2 = 1$ and

$$D_{1,k} = \int \phi_{1,k}\tilde{\phi}.$$

In fact, if we restrict our attention to $[0, 1]$ it follows that V_1 consists of just two scaling functions, and we get

$$\begin{aligned} \phi &= D_{1,0}\phi_{1,0} + D_{1,1}\phi_{1,1} \\ &= D_{1,0}\left[2^{1/2}\phi(2x)\right] + D_{1,1}\left[2^{1/2}\phi(2x+1)\right]. \end{aligned} \tag{2.36}$$

It is easy to see that $D_{1,0} = D_{1,1} = 2^{1/2}$. Extending this to all j we see that $D_{j,k} = 2^{j/2}$. Equation (2.34) implies that the approximation spaces at level $j+1$ yield the wavelet

spaces at level j [31]. Indeed, if $\psi \in W_0$, then

$$\begin{aligned}\psi(x) &= \sum_k (-1)^k D_{1,k}\phi_{1,k}(x) \\ &= \sqrt{2}\,\phi_{1,0}(x) - \sqrt{2}\,\phi_{1,1}(x) \\ &= \sqrt{2}\,\phi(2x) - \sqrt{2}\,\phi(2x-1). \end{aligned} \tag{2.37}$$

The function ψ is the Haar mother wavelet; it is synonymous with the Haar function. In Figure 2.12 the 0th level Haar scaling function $\phi_{0,0}$ is shown (solid line). The dashed line shows the Haar mother wavelet as a linear combination of the two scaling functions at level $j = 1$, normalized in amplitude by $2^{1/2}$ as in (2.37).

From our perspective, the most important aspect of the construction of a wavelet from the scaling functions is that we can now write the wavelet coefficients, $C_{j,k}$, as linear combinations of the scaling coefficients $D_{j,k}$ [31]. We have the following identity, which is a cornerstone of the multiresolution analysis,

$$C_{j,k} = \frac{1}{\sqrt{2}}(D_{j+1,2k} - D_{j+1,2k+1}), \tag{2.38}$$

where

$$C_{j,k} = \int f\psi_{j,k} \tag{2.39a}$$

$$D_{j,k} = \int f\phi_{j,k}. \tag{2.39b}$$

We will make use of these facts in Section 3.5.

2.4.2 Fractal Analysis With Wavelets

Consider a real-valued function $f(x)$. To illustrate the particular way in which wavelets connect to multifractal analysis, suppose that f is $n+1$-times continuously differentiable on the open interval $I = (x_0 - \varepsilon, x_0 + \varepsilon)$, for $\varepsilon > 0$. Then for every

$x \in I$, Taylor's Theorem allows us to expand f as a Taylor series,

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n + R_{n,x_0}(x),$$

where

$$R_{n,x_0}(x) = \int_{x_0}^{x} \frac{f^{n+1}}{(n+1)!}(u - x_0)^{n+1} du.$$

It often happens that the polynomial approximation of f fails at a particular non-integral exponent. Suppose that at x_0 we can best approximate the behavior of f in the neighborhood $|x - x_0|$ by

$$f(x) \approx a_0 + a_1(x - x_0) + \cdots + a_n(x - x_0)^n + c(x - x_0)^{h_n}, \tag{2.40}$$

where $n < h_n < n + 1$. Such "singularities" h_n, or Hölder exponents, can be

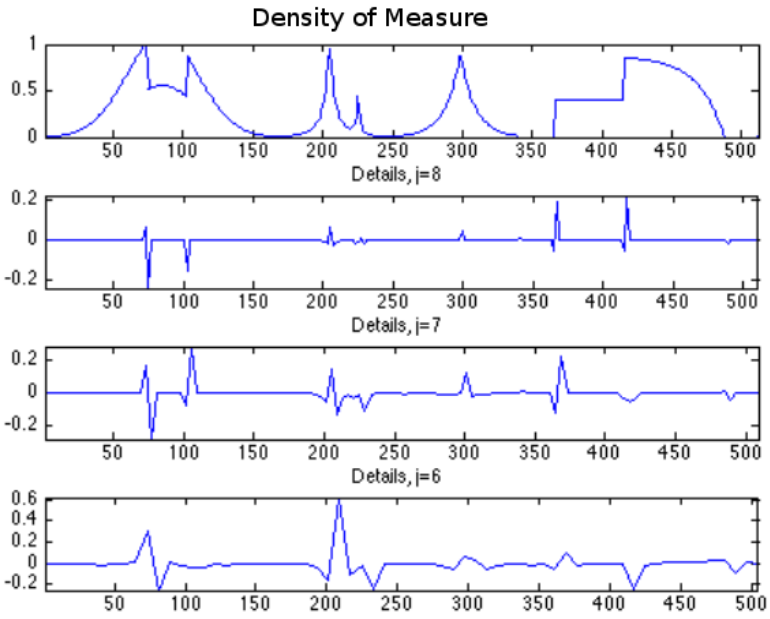


Figure 2.13: The top figure shows a function f with a variety of singularities. The three lower figures that show the wavelet transform of f in order of decreasing resolution.

characterized by the position in space at which they occur as well as the resolution

$|x-x_0|$, since the singularities will appear stronger at different resolutions. Intuitively, if we are able to "kill off" all polynomial behavior of an order less than h_n, we can uncover the particular singularity of order h_n. This is the point at which f becomes a poorly behaved function. In terms of differentiability, if f is n times differentiable, the Hölder exponent h_n indicates that f will fail to be $n+1$ times differentiable. Specifically, in (2.40) the largest integral exponent n such that $h_n \in (n, n+1)$ is the cutoff for "regular" behavior. It is the ability to detect singularities such as Hölder exponents that led to the use of wavelets for fractal analysis of time series [4].

Definition 2.13. Given a function g, suppose that there exists an n-degree polynomial P_n such that

$$|g(x) - P_n(x - x_0)| \leq C|x - x_0|^h \tag{2.41}$$

for all x in an open neighborhood of x_0. Then the supremum of all h such that (2.41) holds is termed the (local) *Hölder exponent* at x_0. We denote the supremum by $h(x_0)$.

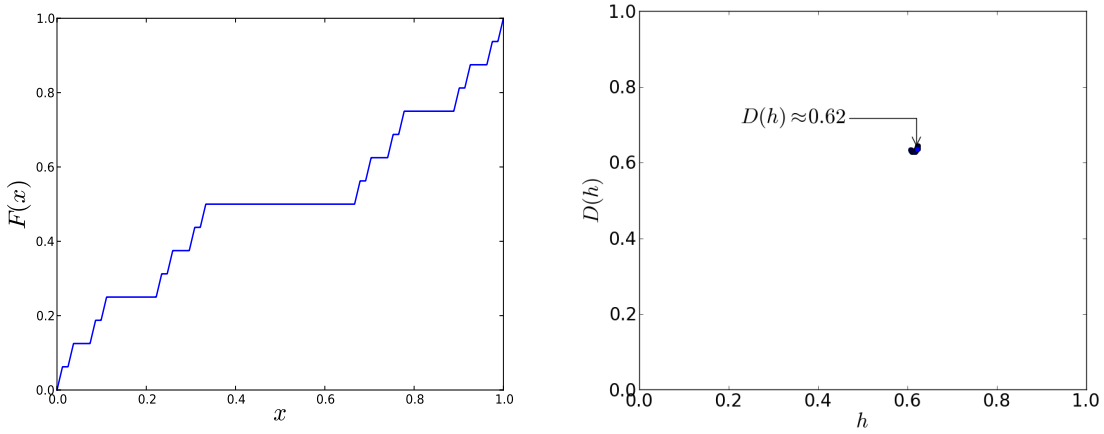
If a wavelet ψ is chosen to be orthogonal to polynomials of order n, that is,

$$\int_{\mathbb{R}} x^n \psi(x) dx = 0, \tag{2.42}$$

we say that ψ has n *vanishing moments*. Assuming condition 2.42 for ψ, then the wavelet transform of a signal f will scale as

$$W_\psi f(a, x) \sim a^{h(x)} \tag{2.43}$$

as $a \to 0^+$ [2, 31]. Figure 2.13 provides an example of how the wavelet transform differs over a small range of scales. A sample function f is plotted in the top row. The wavelet transform of f has been computed at different scales ($a = 2^{-j}$) in the bottom three graphs. The expression in (2.43) can be thought of as computing the density



(a) Devil's Staircase

(b) $D(h)$ Spectrum for (a)

Figure 2.14: (a) The Devil's Staircase is the cumulative distribution of a measure concentrated on the Cantor set. (b) Numerical estimation of the $D(h)$ spectrum using wavelets.

of points in the wavelet within the wavelet's "window". Thus, $W_\psi f(a, x)$ acts as a variable-resolution microscope for f. Figure 2.13 shows the varying density witnessed by $W_\psi f(a, x)$ at three resolutions.

In a time series there may be many points x that have the same Hölder exponent h. Define

$$D(h) := \dim_H \{x \mid h(x) = h\}$$

to be the Hausdorff dimension of the set where the Hölder exponent is h. This is reminiscent of the scaling exponent in the definitions of the Hausdorff dimension in Equation 2.21. Arneodo, *et al.*, showed in a series of papers [2, 5, 10, 62] that the wavelet transform 2.43 can be used to study the singularity spectrum $D(h)$. The wavelet transform acts as a measure at scale a. Thus, the partition function becomes

$$Z(q, a) = \int |W_\psi f(x, a)|^q dx. \tag{2.44}$$

In practice, discrete wavelets [31] are used, in which case Z becomes a sum over discrete points in the domain of f. Analogously to Section 2.3.2, the partition function is used to characterize the fractal properties of a fractal signal through β or by implicitly computing $D(h)$ as in 2.3.5.2 [2, 62].

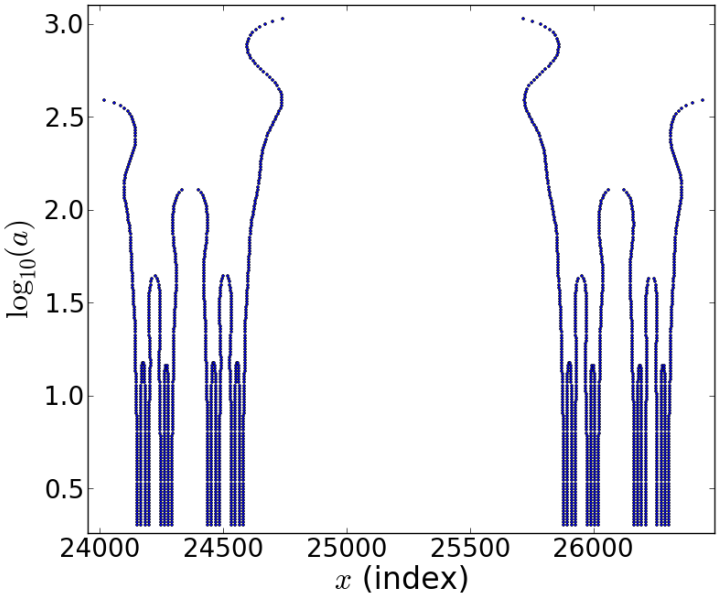


Figure 2.15: Maxima lines for a portion of the space-scale half-plane for the Devil's Staircase. The dyadic structure of the underlying Cantor set is clearly visible.

Arneodo, *et al.*, developed an important variation on this technique. It has allowed the multifractal wavelet analysis to effectively compute the singularity spectrum using negative values of q in the partition function. Historically, negative q have been plagued by divergence issues when balls of small measure, or wavelet transforms of small moduli, are encountered. In practice, this can occur when computing the box-counting dimension and a grid element clips the edge of the set of points. To circumvent this problem, Arneodo, *et al.*, consider an adaptive scheme. Consider the scale a. Instead of computing the partition function for all $|W_\psi f(x,a)|$ at a, choose the wavelet transform $|W_\psi f(x',a')|$ of maximum modulus along a collection

of *maxima lines* descending in scale from the location (x, a). We follow Muzy [62] in defining the set of maxima lines at scale a.

Definition 2.14. The set $L(a)$ consists of all connected curves ℓ satisfying the following:

i. $(x, a') \in \ell \implies a' \leq a$;

ii. for every $a' \leq a$, there exists $(x, a') \in \ell$.

For a certain class of functions resulting from multiplicative processes, Muzy, *et al.* [62] showed that the modified partition function

$$Z(q, a) = \sum_{\ell \in L(a)} \left(\sup_{(x,a') \in \ell} \{|W_\psi f(x, a')|\} \right)^q,$$

behaves the same in the limit as (2.44). This is known as the *wavelet transform modulus maxima* method (WTMM). Note that the reuse of Z above is an abuse of notation that is standard practice, even though it conflicts with the original definition of the partition function in (2.22) and (2.44). This modified Z maintains the desirable scaling property [62]:

$$Z(q, a) \sim a^{-\beta(q)},$$

enabling its use in multifractal analysis. A portion of the set of maxima lines for the Devil's Staircase in Figure 2.14(a) are displayed in Figure 2.15. This type of plot is sometimes referred to as a *wavelet skeleton*. The local maxima at each scale a are followed (connected) to the local maximum or maxima at scale $a' < a$. The maxima lines can be seen to follow the dyadic structure of the underlying Cantor set measure in Figure 2.15.

The simplicial measure developed in Chapter 3 also possesses a localized, adaptive scaling property. In Section 3.5 we show that the simplicial measure on an attractor

can be aligned with the Haar wavelet transform. Hence, by connecting the simplicial measure to the mature field of wavelet analysis, we can utilize the maxima lines approach to study the dimension spectrum.

We present an example showing the agreement between wavelet analysis of multifractal measures on sets and the more traditional analysis in Section 2.3.2.

CHAPTER 3

SIMPLICIAL MEASURES

We develop the concept of a simplicial measure. The primary goal is to produce a cover of a fractal set that is sensitive to the distribution of points. Recall that we seek to approximate the fractal or multifractal measure of the attractor of a dynamical system by studying the way in which a unit mass is distributed over the attractor. In Section 3.1 we describe some basic topological structures and related homology theory that we utilize in our development further on. We narrow our focus to a particular topological construction on a metric space, the Vietoris-Rips (VR) complex [47]. The simplices contained in VR complexes are well suited to our task of computing the mass distribution of points in n-dimensional space as they are constructed in a local manner by considering the mutual distances between points. This construction leads directly to an approximation of the local density of mass in the attractor, and when performed at different resolutions allows us to approximate the multifractal properties of the "natural measure" on an attractor. In Section 3.4, we show that the VR complex on a data set is equivalent to the ideal, minimal covering of a Cantor set by balls of a certain radius (in a manner that we make precise below for finite data points).

3.1 Simplices and Simplicial Complexes

A *simplex* is a geometric object that generalizes the notion of a triangle. A triangle is formed from three vertices all of which are connected to form a 2-dimensional convex hull. In general, an n-dimensional simplex is a polytope formed from the convex hull of $n+1$ vertices. A simplex σ can be denoted by its ordered set of vertices $[v_0, v_1, \ldots, v_n]$

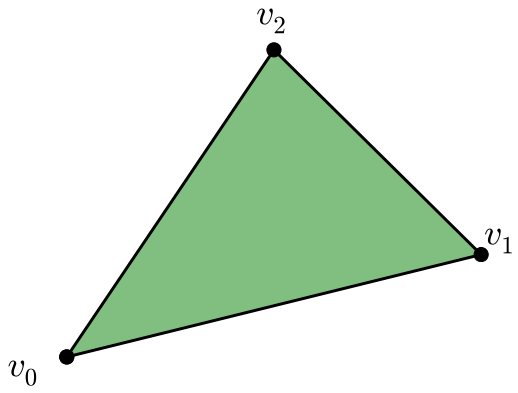


Figure 3.1: A triangle is a 2-dimensional simplex represented by the oriented set of vertices $[v_0, v_1, v_2]$.

or simply $v_0 v_1 \dots v_n$ for brevity [60]. The dimension of σ we denote by $\dim(\sigma)$. For instance, in Figure 3.1, the 2-simplex is defined by the three vertices $v_0, v_1,$ and v_2. A *face* τ of an n-dimensional simplex σ is also a simplex and is defined as the convex hull formed by a (non-empty) subset of the $n+1$ vertices of σ. In Figure 3.1, the 1-simplices $[v_0, v_1]$, $[v_1, v_2]$, $[v_2, v_0]$ are faces along with the three vertices. A *simplicial complex* is a collection Δ of simplices such that for each $\sigma_1, \sigma_2 \in K$, $\sigma_1 \cap \sigma_2$ is a face of both or empty; and any face of a simplex in Δ is also a simplex in Δ. The complement of a face in a given simplex is known as a *coface*. A simplex σ is *maximal* if it has no proper coface in Δ. That is, σ is maximal if it is a face of itself and no other simplex in Δ.

Lastly, a k-skeleton of a simplicial complex Δ is the subcomplex of Δ having faces of dimension no larger than k. In particular, the 1-skeleton of Δ, the collection of 0- and 1-simplices, form the vertices and edges of a graph. Recall that a *clique* is a set of nodes in a graph that include a complete subgraph. If a clique cannot be made any larger then it is termed *maximal*.

3.2 The Vietoris-Rips Complex

We now describe the *Vietoris-Rips complex* (VR complex). The setting is $\mathbb{R}^m$ with a given metric d. We begin by defining the neighborhood graph.

Definition 3.1. Given a finite set of points $Y \subset \mathbb{R}^m$ and a real value $\varepsilon > 0$, the *neighborhood graph* is the undirected graph defined by

$$\mathcal{N}_\varepsilon(Y) := \langle V, E \rangle,$$

where $V = Y$ and $E = \{(x, y) \in Y \times Y \mid d(x, y) < \varepsilon,\ x \neq y\}$.

We work with the Euclidean metric in what follows.* Note that $\mathcal{N}_\varepsilon(Y)$ can be viewed as the 1-skeleton of a simplicial complex. Hence, we can expand $\mathcal{N}_\varepsilon(Y)$ by adding higher dimensional simplices simplices to the set of vertices and edges. We append to V and E a potential simplex σ whenever all edges of σ, $e(\sigma)$ are in E. This is known as a Vietoris-Rips complex [21, 47].

Definition 3.2. Let Y be a finite subset of $\mathbb{R}^m$ and $\varepsilon > 0$. The *Vietoris-Rips complex* of Y at scale ε is the abstract simplicial complex consisting of the vertex set Y, and simplices such that a subset $\{v_0, \ldots, v_k\} \subset Y$ is a simplex iff $\operatorname{diam}(\sigma) < \varepsilon$, where $\operatorname{diam}(\sigma) = \max_{v_i, v_j \in \sigma} \{d(v_i, v_j)\}$. We denote this by

$$\mathcal{R}_\varepsilon(Y) := V \cup E \cup \{\sigma \mid \operatorname{diam}(\sigma) < \varepsilon\}. \tag{3.1}$$

Notice that the maximal cliques in a neighborhood graph are exactly the maximal simplices of the VR complex. The VR complex possesses other useful properties,

*Any metric would suffice, though in general the neighborhood graphs formed from two different metrics are different. In what follows this would yield a different approximation of of the density of points, and ultimately a different multifractal spectrum.

which we shall detail shortly. However, the VR complex has a close topological relative. The construction of the *Čech complex*, is based on intersections, similar to a VR complex [32, 47]. We first state the general definition, then consider a modified definition useful for finite data sets.

Definition 3.3. Let $X \subset \mathbb{R}^m$ and let $\mathcal{U} = \{U_\alpha\}$ be an open covering of X. A Čech complex on X is an abstract simplicial complex in which

i. Each open set in $\mathcal{U}$ is a vertex;

ii. A subcollection $\{U_{\alpha_0}, \ldots, U_{\alpha_k}\}$ of $k+1$ vertices spans a k-simplex iff

$$\bigcap_{i=0}^{k} U_{\alpha_i} \neq \emptyset.$$

Our goals are of a measure-theoretic nature within a metric space, hence it will be useful to formulate Čech complexes in terms of balls of radius ε, which we denote by $B(x, \varepsilon)$ for $x \in \mathbb{R}^m$.

Definition 3.4. Let Y be a finite point set in $\mathbb{R}^m$ and $\varepsilon > 0$. Let

$$\mathcal{B} = \{B(v, \varepsilon) \mid v \in Y\}$$

be an open cover of Y. The Čech complex on Y utilizing ε-balls in $\mathcal{B}$ is the abstract simplicial complex for which

i. Each $v \in Y$ is a vertex;

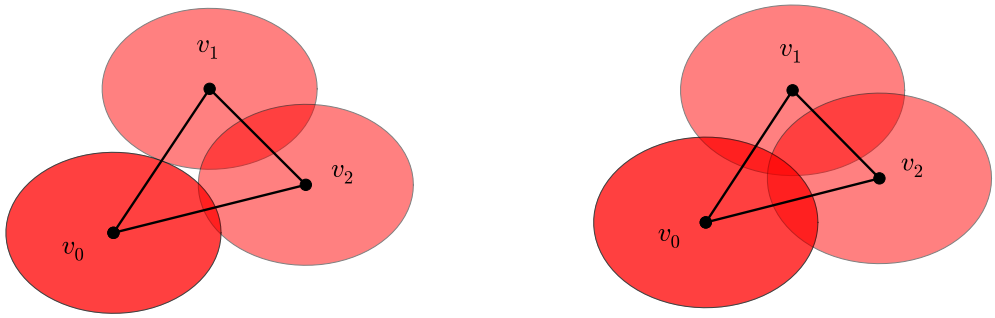
ii. A subcollection of $k+1$ vertices spans a k-simplex iff

$$\bigcap_{i=0}^{k} B(v_i, \varepsilon) \neq \emptyset. \tag{3.2}$$

We indicate a Čech complex of the cover of Y by ε-balls by $\mathcal{C}_\varepsilon(Y)$.

Define a cover of a finite set Y to be the union over $x \in Y$ of balls of radius ε centered at x. If a Čech complex is defined using balls of radius ε, then we will denote the corresponding cover a Čech cover. Topologically, cech covers are quite nice. Define a *good* cover $\mathcal{U}$ to be one for which all sets and all finite intersection of sets in $\mathcal{U}$ are contractible. Then it happens that the Čech complex of a good cover has the same homotopy type as the union of the cover sets [45, 47].

Computationally, the necessity of confirming the nonempty intersection in Equation 3.2 above is highly nontrivial. Hence, the VR complex has proved to be a critical tool for computational topologists as it can be used, in conjunction with Lemma 3.5, to represent the topology of a point set and is relatively fast to compute [77]. This computational tractability is one of the major factors motivating our use of the VR complex in the study of multifractal spectra.



(a) (b)

Figure 3.2: (a) The balls forming the complex have radius ε. The VR complex contains the 2-simplex $[v_0, v_1, v_2]$ while the Čech would not. (b) The result of considering balls of radius $\frac{2}{\sqrt{3}}\varepsilon$ (see Lemma 3.5). Both VR and Čech now contain the same simplices.

Figure 3.2 shows some of the differences between VR and Čech complexes resulting from identical neighborhood graphs formed using slightly difference radii. In the case

of Figure 3.2(a), the Čech complex is composed of the vertices $\{v_0, v_1, v_2\}$ and the three 1-simplices $[v_0, v_1]$, $[v_1, v_2]$, $[v_2, v_0]$. Since all possible edges of the 2-simplex $\sigma = [v_0, v_1, v_2]$ are in the neighborhood graph, the VR complex in Figure 3.2(a) contains σ as well. In Figure 3.2(b), since

$$B\left(v_0, \frac{2}{\sqrt{3}}\varepsilon\right) \cap B\left(v_1, \frac{2}{\sqrt{3}}\varepsilon\right) \cap B\left(v_2, \frac{2}{\sqrt{3}}\varepsilon\right) \neq \emptyset,$$

both the VR and Čech complexes contain the same simplices.

This highlights the topological differences one encounters when comparing VR and Čech complexes. What would appear to be a 1-cycle in Figure 3.2(a) because of the hole in the center is actually the boundary of a 2-simplex σ in the VR complex. That is, the VR complex views σ as contractible. While this difference can prove problematic, especially from the viewpoint of persistent homology [32, 39], it is less so for us, as we are interested primarily in the relationship between VR and Čech complexes as sets and what the simplices in each tell us about the local mass density of an attractor.

We now state Lemma 3.5. Proved by De Silva and Ghrist in [32], it provides an important relationship between VR complexes and Čech complexes, by squeezing a Čech complex formed from the cover of Y by ε-balls between two VR complexes forming slightly different covers of Y.

Lemma 3.5. *For a finite set of points Y in $\mathbb{R}^d$, the Čech complex $\mathcal{C}_\varepsilon(Y)$ is bounded by VR complexes as follows:*

$$\mathcal{R}_{\varepsilon'}(Y) \subset \mathcal{C}_\varepsilon(Y) \subset \mathcal{R}_\varepsilon(Y) \tag{3.3}$$

whenever $\varepsilon \geq \varepsilon' \sqrt{\frac{2d}{d+1}}$.

We prove results in Theorem 3.9 using Čech complexes from $\mathcal{C}_\varepsilon(Y)$. Yet all computation must be performed using VR complexes. Thus, Lemma 3.5 provides a bridge

between the experimental and theoretical sides. We conclude that given any VR complex $\mathcal{R}_\varepsilon(Y)$ there exists a corresponding simplex in $\mathcal{C}_\varepsilon(Y)$.

For instance, in Figure 3.2(a) the vertices of the VR complex are each less than a distance ε' from one another. The centroid falls inside the open neighborhood determined by the intersecting balls in the Čech complex, as in Figure 3.2(b). Thus, if the radius of the balls in Figure 3.2(a) is ε', then a radius of $\frac{2}{\sqrt{3}}\varepsilon'$ suffices to obtain the Čech complex in Figure 3.2(b)

The centroid of a simplex is unique, therefore, for each simplex $\sigma \in \mathcal{R}_{\varepsilon'}(Y)$ and each fixed $\varepsilon'' \in [\varepsilon', \varepsilon]$ there exists a unique ball centered at $y = \phi(\sigma)$ with radius ε'' that contains all of the vertices in σ. Denote the set of centroids in $\mathcal{C}_\varepsilon(Y)$ by

$$\mathcal{M}_\varepsilon(Y) := \{y \in X \mid \phi(\sigma) = y;\, \sigma \in \mathcal{R}_\varepsilon(Y)\}. \tag{3.4}$$

3.3 Simplicial Measures

Our goal is the characterization of dynamical systems from real world data or mathematical models simulating real world processes. We are limited by the finite nature of computation–both in terms of scale and time. A standard tool to analyze such systems in a measure-theoretic context is the residence measure described below. A special case of this is the simplicial measure defined in Subsection 3.3.2.

3.3.1 Residence Measures

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be an ergodic, measure-preserving map on the measure space (X, Σ, λ), where Σ are the Borel-measurable sets in $\mathbb{R}^n$. Suppose that A is an attractor for f. Let Y be a finite set of iterates of f.

A common method for studying such ergodic dynamical systems within the sphere of measure and dimension theory is to define the *residence (probability) measure* on a set $U \subset X$ similarly to (2.1) by

$$\mu(U) = \lim_{m\to\infty} \frac{1}{m} \#\{k \mid 0 \leq k \leq m-1,\ f^k(x) \in U\} \tag{3.5}$$

for λ a.e. $x \in X$. If we assume that λ is an SBR measure, then for certain f the limit exists in (3.5) for Lebesgue a.e. x in a basin of attraction of positive Lebesgue measure [33]. The measure computes the proportion of time that iterates of f spend in U. In the real-world and modeled systems studied only finite data sets are available, denoted as above by the set $Y = \{x_i\}_{i=1}^N$. We approximate the infinite time residence measure from Equation (3.5) on Y by removing the limit. Thus, for a subset $U \subset Y$, let

$$\nu(U) = \frac{1}{N} \#\{k \mid 1 \leq k \leq N,\ f^k(x) \in U\} \tag{3.6}$$

$$= \frac{1}{N} \#\{x \mid x \in U\} \tag{3.7}$$

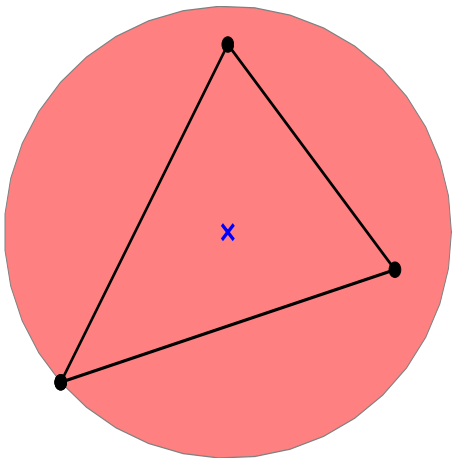


Figure 3.3: A simplex with an associated circumscribing ball B. The centroid is denoted with an $\times$. The ball has a residence measure of $3/N$.

Note that $\nu(Y) = 1$, so it is still a probability measure.

3.3.2 Simplicial Measures

We now define a residence measure based on the simplices in a VR complex. We are interested in using the maximal simplices in a VR complex to understand the density of points in an attractor. First, we abuse notation and define the centroids of $\mathcal{R}_\varepsilon(Y)$ by $\mathcal{M}_\varepsilon(Y)$.

For each $y = \phi(\sigma)$ in $\mathcal{M}_\varepsilon(Y)$, for some $\sigma \in \mathcal{R}_\varepsilon(Y)$, there is a radius $\delta \in [\varepsilon', \varepsilon]$ such that

$$\nu(B(y, \delta)) = \frac{k+1}{N},$$

where $k = \dim(\sigma)$ and N is the cardinality of the data set Y. Therefore, we obtain a probability measure on Y that is determined experimentally by the simplices in a VR complex. We will abuse notation, and talk about the "measure of a simplex" by defining the *simplicial measure* on simplices $\sigma \in \mathcal{R}_\varepsilon(Y)$ as

$$\nu(\sigma) = \nu\left[B(\phi(\sigma), \delta)\right]. \tag{3.8}$$

Remark 3.6. The key objects of study are the maximal simplices in the VR complex $\mathcal{R}_\varepsilon(Y)$. These are formed from the maximal cliques in neighborhood graphs at scale ε. As such, we will neglect much of the algebraic machinery associated to the VR complexes themselves. Nevertheless, we use the notion of maximal simplces as a shorthand based on their construction and equivalence (in the ambient dimension of the data) to maximal cliques.

3.3.3 Trimming Simplicial Complexes

Given a VR complex $\mathcal{R}_\varepsilon(Y)$ and a simplex σ in the complex, $\mathcal{R}_\varepsilon(Y)$ also contains all faces and cofaces of σ. An enumeration of the simplices in $\mathcal{R}_\varepsilon(Y)$ thus contains

far more information than necessary for the purposes of estimating density since all faces of σ contain their own centroids. As such, measures of balls around centroids in $\mathcal{M}_\varepsilon(Y)$ would grossly overestimate the density of a region of space according to the simplicial measure. For example, consider the complex in Figure 3.4. It consists of one 2-simplex (labeled σ in green) and one 3-simplex (also green), as well as a connecting 1-simplex. In addition, each simplex above also contains numerous faces. For instance, σ contains three 0-simplices and three 1-simplices. We would like to avoid enumerating such faces when considering density. Hence, we trim $\mathcal{R}_\varepsilon(Y)$, maintaining only the subset of maximal simplices, $\mathcal{S}_\varepsilon(Y)$. In general, the set of maximal simplices will not be disjoint either. Thus, even with $\mathcal{S}_\varepsilon(Y)$ in hand, we must implement a second round of pruning.

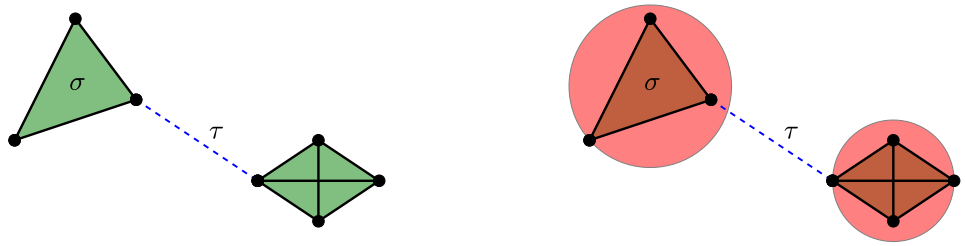
Recall that a collection of sets $\{V_\alpha\}_{\alpha\in\mathcal{J}}$ indexed by $\mathcal{J}$ is said to cover a set Z if

$$\bigcup_\alpha V_\alpha \supset Z. \tag{3.9}$$

The set of maximal simplices $\mathcal{S}_\varepsilon(Y)$ covers Y. These are the maximal cliques in $\mathcal{N}_\varepsilon(Y)$, hence all vertices of Y are indeed covered. We further trim $\mathcal{S}_\varepsilon(Y)$ to produce a *minimal set cover*, which is a subcollection of $\mathcal{S}_\varepsilon(Y)$ of minimal cardinality necessary to cover Y. We term such a cover a *tidy cover*. For the time being we will assume that we are working with a tidy cover of maximal simplices chosen from $\mathcal{R}_\varepsilon(Y)$. We detail the algorithmic process in Section 3.6.2.

3.4 Approximation of Fractal Measures

Having defined the simplicial measure we continue by showing that the approximation of scaling properties of measures on fractal sets is possible using the maximal simplices of VR complexes. We focus on mass distributions and binomial measures.



(a) VR complex (b) Covering with balls of radius $\leq \varepsilon$

Figure 3.4: In this small simplicial complex, the two maximal simplices are the filled (green) 2-simplex and 3-simplex on the left and right, respectively. Since $\dim(\sigma) > \dim(\tau)$, σ will be retained during the trimming process. Similar reasoning causes the 3-simplex on the right to be retained.

We demonstrate that the simplicial measure captures the scaling properties of basic multifractal measures. We connect our results in this context to the robust theory of multifractal wavelet analysis by showing that the simplicial measure and the Haar wavelet coefficients on binomial measures and mass distributions are equivalent.

Recall that an iterated function system (IFS) is an alternative way to construct such measures. As in Equation 2.13, let F be the attractor of an IFS,

$$F = \bigcup_{i=0}^{m-1} S_i(F),$$

where $\{S_0, \ldots, S_{m-1}\}$ is a finite family of contractions on $[0, 1]$. Below we consider approximations to F, where F_n denotes the nth level of the construction of F. Truncating the construction of F at the nth level we get the *prefractal* composed of m^n regions,

$$F_n = \bigcup_{j=0}^{m^n-1} F_j^{(n)}.$$

There is a large body of work dealing with binomial measures and mass distributions described using two contractions in the context of fractals and multifractal

spectra [1, 2, 10, 36, 43, 62, 67]. In order to align the simplicial measure with this work we will consider the class of IFS's constructed with contractions $\{S_0, S_1\}$ such that $S_i : [0,1] \to [0,1]$, with a common contraction ratio c. For the remainder of the section we will assume that $m = 2$. Each contraction is associated with a conserved split in the mass at each level such that at level n

$$p_{k_n}(l) \mapsto \{p_{k_n}(l)p_0, p_{k_n}(l)p_1\},$$

where $k_n(l) \in \{0,1\}^n$, $l \in \{0, 1, \ldots, 2^n - 1\}$, and $p_{k_n(l)} = \prod_{j=1}^{l} p_{i_j}$, with $i_j \in \{0, 1\}$.

Let μ be a mass distribution on F. Given $n \in \mathbb{N}$ and the IFS $\{S_0, S_1\}$ with common contraction ratio c, then $\mu\left(F_j^{(n)}\right) = p_{k_n(j)}$, and each $F_j^{(n)} \in F_n$ is equivalent to an interval of $I_j^{(n)}$ of diameter c^n.

We define the minimum gap between intervals as

$$\gamma = \gamma_n := \min_{j=0,\ldots,m^n-2} d\left(F_j^{(n)}, F_{j+1}^{(n)}\right),$$

where $d(A, B) = \inf_{x \in A} d(x, B)$. (Strictly speaking, this is more general than necessary since with contractions S_0 and S_1 each gap is of the same size at each level n.)

The simplicial measure deals with finite data sets. Given $n \in \mathbb{N}$, let $Y_n = \{x_1, \ldots, x_N\} \subset F_n$, where N is chosen so that for each $j \in \{0, \ldots, m^n - 1\}$

$$\left|\mu(F_j^{(n)}) - \mu_N(F_j^{(n)})\right| \leq \frac{1}{N}, \tag{3.10}$$

where

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i},$$

with $\delta_x(A)$ defined in Equation (2.1). In particular, we assume that the points in Y faithfully approximate μ so that as $N \to \infty$,

$$\mu_N \to \mu. \tag{3.11}$$

In addition, assume that for each j the points in Y_n are distributed uniformly across $F_j^{(n)}$ according to the measure $p_{k_n}(j)$. If Y_n is a sequence of points in the trajectory of a dynamical system, then we say that the ergodic averages converge to μ whenever (3.11) holds. Condition 3.10 is a fairly natural assumption, and occurs *by construction* when one experimentally approximates the attractor of an IFS.

In general, simplices may exist in a higher dimension than Y_n. We define a projection from this higher-dimensional space to the ambient space in which Y_n resides.

Definition 3.7. Suppose $Y_n \subset \mathbb{R}^b$ is given, and $\sigma \in \mathbb{R}^{\dim \sigma}$ is a simplex formed from vertices in Y_n. (I.e., σ is in a Čech or VR complex formed on the points in Y_n.) Consider a point $x \in \sigma$, then $x = \sum_{i=0}^{\dim \sigma} t_i v_i$, where $v_i \in \mathbb{R}^{\dim \sigma}$ are the vertices of σ and $\sum_{i=0}^{\dim \sigma} t_i = 1$. Define the projection map back to Y_n by $\pi : \mathbb{R}^{\dim \sigma} \to \mathbb{R}^b$ by

$$\pi(x, Y_n) = \pi(x) = \sum_{i=0}^{\dim \sigma} t_i \pi(v_i) = \sum_{i=0}^{\dim \sigma} t_i u_i,$$

where u_i is the corresponding vertex in Y_n. By $\pi(\sigma)$ we will also denote the set that is the projection of all $x \in \sigma$ into Y_n.

Recall that $\mathcal{S}_\varepsilon(Y_n)$ denotes the trimmed set of maximal simplices in $\mathcal{R}_\varepsilon(Y_n)$.

Definition 3.8. Let E be a Borel set in $\mathbb{R}^b$, where b is the dimension of the data in Y_n. Suppose Y_n is given for $n \in \mathbb{N}$, and let $\varepsilon > 0$. Then we define

$$\nu_n(E) = \sum_{\sigma \in \mathcal{S}_\varepsilon(Y_n) : \pi(\sigma) \subset E} \nu(\sigma), \tag{3.12}$$

where ν is the simplicial measure associated to Y_n and ε.

We now state our main result. We examine the case when the radius of balls about points in Y_n allows a straightforward division of the intervals in F_n. For such a case, there are two requirements to create a disjoint cover of the points in $Y_n \subset F_n$. First,

the radius chosen to form the VR complex must be small enough to avoid overlap with maximal simplices from neighboring intervals in F. Second, the simplices must evenly divide the intervals. We describe this below.

Theorem 3.9. *For all $n \in \mathbb{N}$, suppose $N \in \mathbb{N}$ and $Y_n \subset F_n$ are chosen so that* (3.10) *holds. Let $r, s \in \mathbb{N}$ and $\alpha = \frac{1}{r^s}$ be such that $\varepsilon = \varepsilon_n = \alpha c^n < \frac{\gamma}{2}$. If r is divisible by 2, then there exists a disjoint cover of Y_n of maximal simplices in $\mathcal{S}_\varepsilon(Y_n)$ such that $\forall F_j^{(n)} \in F_n$,*

$$\nu_n\left(F_j^{(n)}\right) \to \mu\left(F_j^{(n)}\right).$$

Remark 3.10. The choice of ε guarantees that the cover of each $F_j^{(n)}$ is disjoint from $F_{j'}^{(n)}$, for $j \neq j' \in \{0\ldots., m^n - 1\}$.

We prove Theorem 3.9 in the technical lemmas that follow. Given Y_n with cardinality N, let $Y_j^{(n)} = \left(Y_n \cap F_j^{(n)}\right)$ and $m_j = m_j^{(n)} = \#Y_j^{(n)}$.

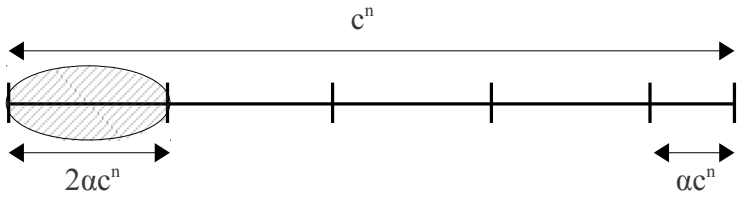


Figure 3.5: Example of a covering of $F_j^{(n)}$ by disjoint maximal cliques (simplices in $\mathcal{S}_\varepsilon(Y_n)$) of diameter $2\alpha c^n$. The case where r is not divisible by 2 leaves a portion on the right-hand side uncovered my maximal simplices.

Lemma 3.11. *Suppose n, N, Y_n, and α are as in Theorem 3.9. Then there exists a subcollection of disjoint maximal cliques in $\mathcal{S}_\varepsilon(Y_n)$ whose projected vertices cover Y_n.*

Proof. Consider the subcollection of maximal simplices in $\mathcal{S}_\varepsilon(Y_n)$ whose vertices are projected onto $Y_n \cap F_j^{(n)}$, which we define by $C_j^{(n)} = \{\sigma \in \mathcal{S}_\varepsilon(Y_n) \mid \pi(\sigma) \cap Y_j^{(n)} \neq \emptyset\}$. Choose $\sigma_0 \in C_j^{(n)}$ for which there exists $x \in \pi(\sigma_0)$ such that $x \leq y$ for all $y \in Y_j^{(n)}$.

That is, choose the left-most simplex in $C_j^{(n)}$. Form the collection $A_j^{(n)} = \{\sigma_0\}$. We will collect disjoint maximal simplices in $A_j^{(n)}$ to obtain a cover of $Y_j^{(n)}$. Consider the subset of $C_j^{(n)}$ of simplices disjoint from that in $A_j^{(n)}$,

$$D = D(A_j^{(n)}) = \left\{\sigma \in C_j^{(n)} \mid \sigma \cap \tau = \emptyset, \forall \tau \in A_j^{(n)}\right\}.$$

In the same manner as above, choose σ_1 in D to be the left-most element in D and append σ_1 to $A_j^{(n)}$. Repeat this process until no more simplices can be added disjointly, obtaining the collection $A_j^{(n)} = \{\sigma_0, \ldots, \sigma_t\}$. It remains to determine t and to show that the simplices in $A_j^{(n)}$ cover $Y_j^{(n)}$.

Each $\sigma \in A_j^{(n)}$ has a diameter in Y_n of

$$|\pi(\sigma)| = 2\alpha c^n = 2 \cdot (2q)^{-s} c^n = \frac{c^n}{2^{s-1} q^s},$$

where $r = 2q$, by the assumption on the divisibility of r. We can partition $F_j^{(n)}$ into $t = 2^{1-s} q^s$ intervals of length $|\pi(\sigma)|$. By construction, for each $i \neq i'$, the vertices of $\pi(\sigma_i)$ are disjoint from $\pi(\sigma_{i'})$. Hence, $A_j^{(n)}$ forms a corresponding disjoint partition of the points in $Y_j^{(n)}$. The choice of α gaurantees that the cover of $Y_j^{(n)}$ is disjoint from other $Y_{j'}^{(n)}$. Thus,

$$\bigcup_{j=0}^{2^n-1} A_j^{(n)}$$

covers Y_n disjointly. □

We examine the case where r is not divisible by two. This is direct results of the geometry of the vertices of simplices in the intervals of F_n.

Lemma 3.12. *Let $n, N, Y_j^{(n)}, F_j^{(n)}$ be as above. If $r \in \mathbb{N}$, $\alpha = r^{-s}$ and r is not divisible by 2, then a disjoint collection $A_j^{(n)}$ of maximal simplices in $\mathcal{S}_\varepsilon(Y_n)$ will not cover $Y_j^{(n)}$. The number of points not covered by simplices in $A_j^{(n)}$ is αm_j.*

Proof. Suppose that $A_j^{(n)}$ is a disjoint collection of simplices constructed as in Lemma 3.11. For each $\sigma \in A_j^{(n)}$,

$$|\pi(\sigma)| = 2\left(\frac{c^n}{r^s}\right).$$

Since $\frac{r}{2}$ is not integral, division of $F_j^{(n)}$ from left to right into intervals of length $2\left(\frac{c^n}{r^s}\right)$ results in an extra interval E_α of length $\frac{c^n}{r^s}$ on the right side of $F_j^{(n)}$. Each $\sigma \in A_j^{(n)}$ has a projected diameter of $2\left(\frac{c^n}{r^s}\right)$. Thus, the construction of $A_j^{(n)}$ will not cover this final segment of length $\frac{c^n}{r^s}$, since no maximal simplex in $\mathcal{S}_\varepsilon(Y_n)$ can cover E_α without intersecting an element of $A_j^{(n)}$. Given the uniform distribution of points in $Y_j^{(n)}$, the simplices in $A_j^{(n)}$ will cover $(1-\alpha)m_j$ points in $Y_j^{(n)}$, leaving αm_j uncovered. □

Proof of Theorem 3.9 Let $j \in \{0, \ldots, 2^n - 1\}$ and set $E = F_j^{(n)}$. Then by (3.12)

$$\nu_n(E) = \mu_N(E) = \frac{\#(Y_n \cap F_j^{(n)})}{N}. \tag{3.13}$$

Hence, by (3.10), $\nu_n(E) \to \mu(E)$ as $N \to \infty$. Theorem 3.9 follows. □

3.5 Equivalence of Simplicial Measures and Wavelet Methods

For finite data the construction of maximal simplices and the corresponding simplicial measure allows for the approximation of the local intensity of a measure. We turn now to an analysis of certain measures using wavelet theory. The wavelet transform multifractal formalism [10] is a mature theory developed for the application of wavelets to singular measures such as those studied in Section 3.4 and fractal signals. We show that the simplicial measure can be viewed as an approximation of the Haar wavelet transform on such singular measures, thereby linking the two techniques.

3.5.1 The Setting

Let μ be a self-similar mass distribution or binomial probability measure as in Section 3.4. Define the dyadic intervals $I^n_{n_k} := F_{n_k}$ to be those used in the construction of μ as above as well, where $n_k \in \{0, 1, \ldots, 2^n - 1\}$. Assume that

$$\mu([0,1]) = \mu(I^0_0) = p^0_0 = 1.$$

Recall from Section 2.3.3 on page 17 that

$$\mu(I^n_{n_k}) = p^0_0 p^1_{k_1} \cdots p^n_{n_k}, \tag{3.14}$$

where we retain p^0_0 for consistency. Let g be a measurable function. We write the integral of g with respect to the measure μ as

$$\lim_{n\to\infty} \sum_{j=0}^{2^n-1} g(j2^{-n})\mu(I^n_j) = \int g(t)\mu(dt). \tag{3.15}$$

3.5.2 The Haar Wavelet

Recall from Section 2.4 that a multiresolution analysis consists of sequences of discrete scaling and wavelet functions. Together these functions partition $L^2(\mathbb{R})$ into orthogonal scaling and wavelet subspaces that fit together in a nested, dyadic manner.

The following relationship between the scaling and wavelet transforms from the multiresolution analysis in Subsection 2.4.1 holds. For

$$C_{j,k} = \int f\psi_{j,k} \tag{3.16a}$$

and

$$D_{j,k} = \int f\phi_{j,k}, \tag{3.16b}$$

with $\phi_{j,k} \in V_j$ and $\psi_{j,k} \in W_j$, we have that

$$C_{j,k} = \frac{1}{\sqrt{2}}(D_{j+1,2k} - D_{j+1,2k+1}). \tag{3.17}$$

Thus, for certain fractal measures, Equation 3.17 allows us to connect the wavelet transform with the simplicial measure.

3.5.3 Wavelet Analysis of μ

Mass distributions and binomial measures in Sections 2.3.3 are defined on very regular partitions of the unit interval. Such constructions are ideally suited to a multiresolution analysis using Haar scaling and wavelet functions. We consider the binomial measure μ and analyze it using the scaling functions in order to derive appropriate wavelet coefficients. In the wavelet transform multifractal formalism, these are used to obtain the multifractal spectrum of μ. (See Section 2.4 and [62].)

Using (3.14) and (3.15), along with the fact that each $\phi_{j,k}$ is a suitably-scaled indicator function, we write that the scaling coefficients with respect to the measure μ as

$$\begin{aligned} D_{j,k_j} &= \int \phi_{j,k_j}\mu(dt) \\ &= 2^{j/2} \int_{k_j 2^{-j}}^{(k_j+1)2^{-j}} \mu(dt) \\ &= 2^{j/2}\mu\left(I_{k_j}^j\right). \end{aligned} \tag{3.18}$$

Thus, using (3.17) and (3.18), we can rewrite the wavelet coefficients in terms of the measure,

$$C_{j,k} = \frac{1}{\sqrt{2}}\left(2^{j/2}\mu\left(I_{2k}^{j+1}\right) - 2^{j/2}\mu\left(I_{2k+1}^{j+1}\right)\right) \tag{3.19}$$

Rearranging (3.19) gives the wavelet coefficients at scale j entirely in terms of the measures of intervals at scale $j+1$,

$$C_{j,k}\sqrt{2^{(1-j)}} = \left(\mu\left(I_{2k}^{j+1}\right) - \mu\left(I_{2k+1}^{j+1}\right)\right). \tag{3.20}$$

From Section 3.4 we know that the simplicial measure ν approximates μ. Hence, it follows that at each resolution the Haar wavelet can be approximated by ν if we replace μ with ν in 3.20. Furthermore, by Theorem 3.9, given an arbitrarily large number of data points, we can replace we can replace μ with ν in (3.20).

3.6 Algorithms

Up to this point we have not discussed the way in which the simplices used for the simplicial measures are constructed. The task of computing the multifractal spectrum from a finite number of iterates of a dynamical system is, at its heart, an approximation problem. Given a finite set of iterates, $Y = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$, of a dynamical system f, we seek to group these into sets at different scales in such as way as to estimate the density of points in the attractor of f. A number of methods have been developed to do this. As described in Section 3.3, we use the trimmed set of maximal simplices to estimate of the density of points at various resolutions. In the remainder of this section we consider the computational and algorithmic factors involved in this approach.

All of the software developed for this dissertation is located in a central repository at [16].

3.6.1 Neighborhood Graphs and Maximal Simplex Construction

The problem of finding the nearest neighbors of a point set is fundamental in computational geometry and search algorithms in general Arya, *et al.* [9]. Let Y be

a set of n points in d-dimensional space. Given a query point $z \in Y$, its nearest neighbor is the point

$$x^* = \min_{x \neq z} \|z - x\|,$$

where $\|\cdot\|$ is typically Euclidean distance. Extending this problem, we require sets of neighbors within ε of all points in Y. The simplest approach is to compare all pairs of points by constructing an $n \times n$ distance matrix. Of course, the $O(n^2)$ time and space requirements of this approach are prohibitive.

Actually, it is not necessary to compare all pairs, and search algorithms exist that utilize *kd-trees* to organize and partition the data prior to assembling a list of nearest neighbors [24]. A significant improvement in performance is gained, yielding $O(\log n)$ query time and $O(n)$ space requirements. If dimension of the data is relatively high (> 20), then prioritized searches such as kd-trees lose their advantage over brute force queries. An *approximate nearest neighbor* algorithm can be employed [8, 77]. Given the query point z with a true nearest neighbor x^*, an approximate nearest neighbor x is one for which

$$d(z, x) \leq (1 + \delta) d(z, x^*),$$

where $\delta > 0$ is some small, allowable error. We note the existence of the approximate nearest neighbor algorithm to emphasize that the algorithms used to construct VR complexes are scalable to higher dimensions, if one is willing to accept a small amount of error. In experiments comparing exact and approximate distance algorithms, there is no difference seen in the results for multifractal spectra. This is likely because the magnitude of maximal simplices are changed only slightly in the VR complex by the addition of δ, and hence the partition function is also little changed.

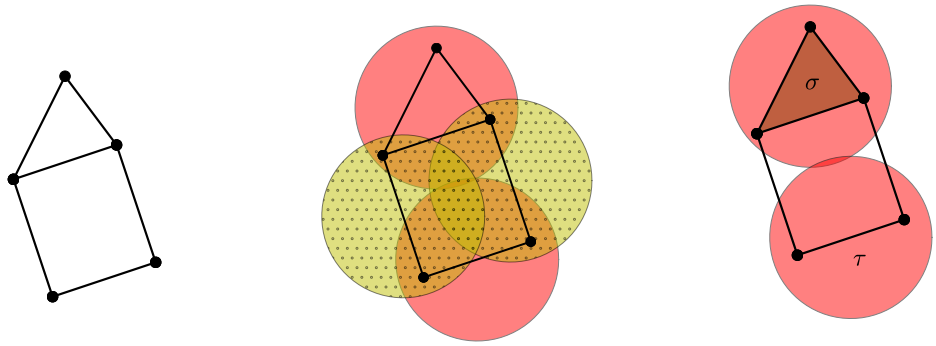
The results reported in Chapter 4 use the exact distance algorithm. Either algorithm stores the output in the form of a hash table [24] keyed by the points in Y.

The ith point, x_i, is mapped to its k nearest neighbors in the table, eg.,

$$x_i \mapsto (y_{i_1}, \ldots, y_{i_k}), \tag{3.21}$$

In our approach, no bound is put on k, since we are interested in the collection of points within $\varepsilon > 0$ of a point in Y. This query also yields a hash table as in 3.21, except that each key x_i now maps to a variable number of neighbors, say N_i. This data structure yields the desired neighborhood graph $\mathcal{N}_\varepsilon(Y) = \langle V, E \rangle$, with nodes $V = Y$, and the edge set E formed from all sets of pairs $\{(x_i, y_{i_j})\}_{j=1}^{N_i}$ in the hash table of neighbors.

Given $\varepsilon > 0$, a finite data set Y, and the neighborhood graph $\mathcal{N}_\varepsilon(Y)$, we know from Section 3.2 that $\mathcal{N}_\varepsilon(Y)$ can be expanded to a VR complex $\mathcal{R}_\varepsilon(Y)$. This is a combinatorial problem that is now independent of the geometry of the underlying space. We focus our attention on a partial expansion, constructing only the maximal simplices of $\mathcal{R}_\varepsilon(Y)$.



(a) Clique complex (b) Simplices in $\mathcal{S}(Y)$ (c) Simplices in $\mathcal{T}(Y)$

Figure 3.6: (a) A neighborhood graph composed of 5 nodes. (b) The circumscribing balls of the maximal simplices in $\mathcal{S}(Y)$. The left and right balls (yellow with hatching) indicate over counting. (c) The subset $\mathcal{T}(Y)$ of maximal simplices after trimming. The circumscribing balls are centered at the centroids of σ and τ.

Let $G = \langle V, E \rangle$ be an undirected graph without self-loops, where $V = V(G)$ is the set of vertices and $E = E(G)$ is the edge set. Recall that a *clique* is a subset of vertices $A \subset V$ such that there is an edge between each pair of vertices $u, v \in A$. A maximal clique is a subset of vertices that cannot be expanded by the addition of one more vertex. Given an arbitrary graph G the *clique complex* is an abstract simplicial complex whose vertex set is $V(G)$ and whose simplices consist of all complete subgraphs in G [53]. The VR complex C for a neighborhood graph G is the maximal clique complex for G. Hence, the maximal simplices in C are exactly the maximal cliques in G.

The algorithm for finding maximal cliques is based on the one described in the classic paper on the subject by Bron and Kerbosch in 1973 [20]. To give an idea of how one constructs maximal cliques from a graph G, Algorithm BK reviews the basic recursive algorithm developed by Bron and Kerbosch. Note that R and X are initialized to the empty set when $\mathrm{BK}(V, \emptyset, \emptyset)$ is called. On recursive calls (line 10) they are passed in non-empty. Neighbors of a vertex u are referred to by $N(u)$. (The algorithm in the form presented is from Cazals and Karnade [22].)

BK(V, R, X)

```
 1  if V = ∅ and X = ∅
 2      Return R as maximal clique
 3  else
 4      Assume V = {u_1, ..., u_n}
 5      for i = 1 to n
 6
 7          R_new = R ∪ u_i
 8          V_new = V ∩ N(u_i)          // Nodes connected to nodes in R
 9          X_new = X ∩ N(u_i)
10          BK(V_new, R_new, X_new)     // Backtrack WRT u_i
11          V = V − {u_i}
12          X = X ∪ {u_i}               // Nodes already processed
```

The three sets in the BK algorithm are disjoint. Once a vertex $u \in V$ has been chosen, then within the recursive call on line 10, V and X are restricted to neighbors of u. The recursion backtracks through these neighbors to guarantee that only mutually-connected vertices are appended to R. Only when this subset of vertices has been exhausted does the algorithm move on to the next vertex in V.

The BK algorithm can only discern that a clique is maximal or non-maximal after all possible cliques have been formed. Therefore, all cliques are formed in BK. Significant reductions to the size of the recursion tree are possible through the introduction of a *pivot* which truncates the tree and avoids repeating subtrees [22].

3.6.2 Trimming the Collection of Maximal Simplices

Let $\mathcal{S}_\varepsilon(Y) \subset \mathcal{R}_\varepsilon(Y)$ be the set of maximal cliques in the neighborhood graph $\mathcal{N}_\varepsilon(Y)$ at resolution ε. The example in Figure 3.6(b) shows the drawback inherent in

$\mathcal{S}_\varepsilon(Y)$ with respect to an accurate analysis of an attractor using a data set Y. When $\mathcal{N}_\varepsilon(Y)$ is expanded to a VR complex there are five 1-vertex cliques, six 2-vertex cliques, and three 3-vertex clique. The 3-vertex clique is maximal as are the three 2-vertex cliques that include the two lower vertices. The vertices in the 1-simplices on the left and right (yellow, hatched balls) are included in the two simplices highlighted in Figure 3.6(c). The maximal simplices in $\mathcal{S}_\varepsilon(Y)$ are used to measure the number of data points in a ball whose center is located at the centroid of the simplex. As Figure 3.6(b) shows, without any modification the collection of maximal simplices will, in general, measure points residing in different circumscribing balls multiple times. We seek, therefore, to select an optimal subcollection of maximal simplices that accurately approximates the density of points in an attractor. This problem is an example of another classic problem in computer science, the *set cover problem* [24].

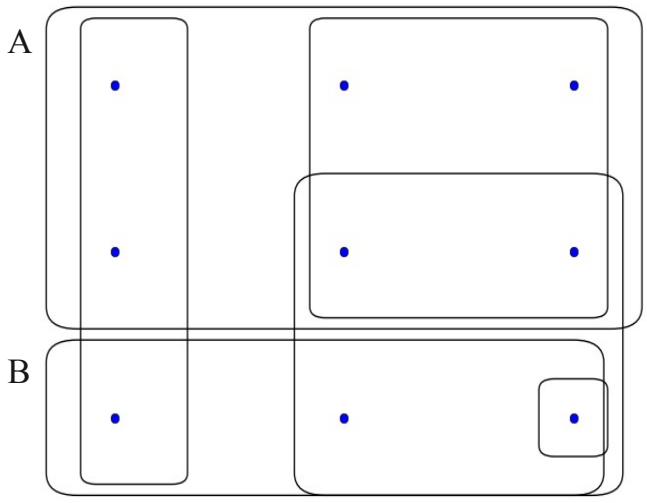


Figure 3.7: A set of nine points with six subsets. The optimal cover contains only the two sets A and B.

The set cover problem is defined as follows [24]: Given a "universe" X and a family S of subsets of X, define a *cover* of X to be

$$C \subset S \text{ such that } X \subset \bigcup_{A \in C} A. \tag{3.22}$$

A *minimal set cover* is a cover $\hat{C}$ such that the cardinality of C, denoted $\#C$, is minimal. The problem of finding the optimal set cover is NP-hard [24]. Figure 3.7 is a basic example of consisting of a universe X of nine points. Each of the six rectangles represents a set in S containing those points falling within its boundary. The minimal cover consists of the subcollection $\{A, B\}$.

In the context of dynamical systems and simplicial measures, the universe is the finite data set Y, the family S is, in general, the set of all possible simplices, and the cover C is a subset of those simplices such that for all $x \in Y$ there exists a simplex $\sigma \in C$ with $x \in [v_0, \ldots, v_n] = \sigma$. If we restrict S to simplices in $\mathcal{S}_\varepsilon(Y)$, then the minimal cover contains the least number of *maximal* simplices necessary to cover Y. We term this the *min-max cover*. We use a greedy heuristic to obtain a "good" minimal set cover using the greedy algorithm. It provides an approximation (in polynomial time) of the true minimal cover. Because the algorithm is *greedy* it will choose the sets $A \in S$ that maximize $|A \cap Y|$. We describe this algorithm below [24]:

SET-COVER(X, S)

1 $U = X$

2 $C = \emptyset$

3 **while** $U \neq \emptyset$

4 Choose $A \in S$ that maximizes $|A \cap X|$

5 $U = U - A$

6 $C = C \cup A$

7 **return** C

Each iteration of the **while** loop greedily chooses the largest set $A \in S$ that has not been considered yet to add to the cover C until all elements of U are covered. On line 5, A is removed from U; and on line 6 A is added to the cover C. The SET-COVER heuristic exits the **while** loop and returns C precisely when Equation 3.22 is satisfied.

The drawback of the SET-COVER heuristic is that nothing prevents it from returning a cover composed of many overlapping sets, as long as those sets are greedily chosen by line 4. Therefore, we implement a variant of the SET-COVER heuristic that replaces line 4 to create a set C' that is a *disjoint* collection of sets:

Choose $A \in S$ that maximizes $|A \cap X|$ and $A \cap C' = \emptyset$

In general, $\bigcup_{A \in C'} A \neq X$, so the above introduces an infinite loop into SET-COVER since, in general, it will never happen that $U = \emptyset$. There are ways to circumvent this problem, which we detail below. In terms of understanding the natural measure on the attractor of a dynamical system, though, C' presents us with the densest *disjoint* regions in the finite approximation.

As mentioned, the **while** loop between lines 3 and 6 may never exit. We consider an alternative way to check whether we have considered all sets in the family of sets S.

DISJOINT-SET-COVER(X, S)

```
1  U = X
2  N = #X                    // Number of points in the "universe"
3  C = ∅
4  K = ∅                     // Unique points checked
5  while S ≠ ∅
6      Choose A ∈ S
7      K = K ∪ A             // Add points in A to points checked
8      if |A ∩ X| and A ∩ C' = ∅
9          S = S − A
10         C = C ∪ A
11     if #K = N:            // Short-circuit if we've checked all points
12         return C
13 return C
```

There is a possible increase in complexity since DISJOINT-SET-COVER may perform more loops in its search since it discards overlapping sets that otherwise would have been chosen by SET-COVER. Yet, in practice DISJOINT-SET-COVER the short-circuit in line 11. Thus, even though DISJOINT-SET-COVER does not cover every point in the universe X, the sets of points that it does (greedily) cover a assured to be the highest-density disjoint regions.

3.6.3 Lines of Local Maxima

The analysis of fractals by wavelet methods using the wavelet transform modulus maxima (WTMM) method has proven remarkably successful since its introduction in the early 1990's. The key idea behind this method is the introduction of a variable resolution "box". In the case of the simplicial measure, these "boxes" are the maximal

simplices at each resolution ε. Considering the suprema along a maxima line enables a heterogeneous covering of the fractal by simplices of radius no greater than ε.

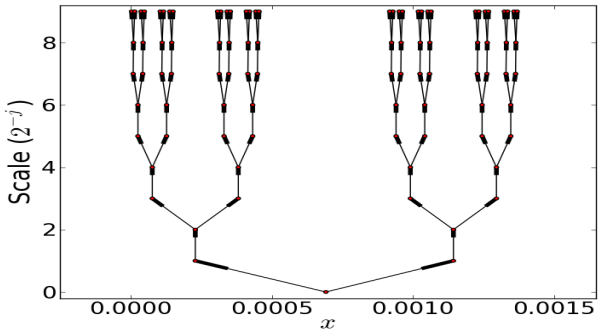


Figure 3.8: Graph structure showing the maxima lines for a Cantor set. The red circles represent the location in the scale-space half-plane of the centroids of maximal simplices. Only a small portion of the maxima lines "skeleton" is shown.

Consider Figure 3.8 showing a portion of the maxima lines "skeleton" constructed using maximal simplices on the Cantor set. In practice, we construct a graph consisting of nodes and edges. The nodes are situated at the centroids of maximal simplices. The nodes are arranged vertically according to scale, in that each horizontal level contains maximal simplices at a fixed resolution. Each horizontal layer of nodes consists of the local maxima of the maximal simplices at that scale. The edges connect each node at scale ε with its nearest neighbors in the spacial direction at the next smaller scale ε'. For instance, in Figure 3.8 we see that this detects when the maxima lines bifurcate at nodes at scale 2^{-4}, with each node connecting to two node at scale 2^{-5}. In this way we create the maxima lines visible in Figure 3.8. The dyadic construction of the Cantor set is clearly visible in graph.

Creating maxima graphs such as Figure 3.8 can be accomplished efficiently in higher dimensions by creating layers of kd-trees [24] whose data is composed of the local maxima of the maximal simplices. The power of this approach is that querying

a node in a kd-tree at scale ε for its nearest neighbors at scale ε' can be accomplished in an algorithmically efficient manner.

The WTMM method works well for a subtle reason. When a wavelet's width is scaled by a factor of 2^{-j}, its amplitude is scaled simultaneously my a factor $2^{j/2}$. This enables the comparison of the wavelet transform of a function across scales, since any loss of measure due to narrowing the "box" is compensated for in the amplitude. We should mention that at this juncture the algorithms necessary to utilize the maxima lines graph for the simplicial measure have proven effective only on the Cantor set. One issue likely stems from the fact that in a structure without an clear multiplicative structure, such as the Hénon attractor, the circumscribing balls for the maximal simplices have a variable footprint. Hence, we find it difficult to compare across scales in as uniform a manner as that afforded by the wavelet transform.

CHAPTER 4
APPLICATIONS

The analysis of dynamical systems using computational means is a non-trivial endeavor. Simple examples that can be solved analytically are necessary to examine the algorithms. In Chapter 3 we developed the computational algorithms necessary to simplicial measure theory. In this chapter we apply the algorithms to a variety of problems, from theoretical to real-world data

In Section 4.1 we show, using the Cantor set, that simplicial measure provide impressive numerical approximations of the multifractal spectrum of attractors. A traditional test piece for numerical methods applied to the study of dynamical systems is the Henon map. Section 4.2 is devoted to a study of the fractal dimension of the Henon map using simplicial measures.

In developing the computational methods in Chapter 3 the ultimate goal is to apply them to data generated by mathematical models as well as data collected from real sources. In Section 4.3, we demonstrate a mathematical model of a neural network, first reported in [17], which generates a time series that shows the characteristics of physiological systems using a couple of well-established metrics. Next, we show in Section 4.3.8 that the multifractal spectrum, computed using the simplicial measure, shows a range for the scaling exponent α close to that exhibited by healthy physiological systems.

Lastly, in Section 4.5 we apply multifractal analysis to heartbeat data from healthy and compromised subjects. In terms of the range of the scaling exponents, our results results closely align with previous results of Ivanov, *et al.* [49].

4.1 Cantor Set Results

Recall that the Cantor set is constructed by intersecting the images of an infinite sequence of iterations of an iterated function system (IFS) acting on the unit interval. The Cantor set can be approximated by a finite sequence of iterations the IFS. Simulating an attractor of a dynamical system which has the properties of a Cantor set is accomplished by populating the finite set of intervals with various distributions of points. By changing the distributions we obtain a spectrum of multifractal spectra, an example of which is seen in Figure 4.3.

4.1.1 Approximation of the Attractor

We begin with a unit mass uniformly distributed across the unit interval $E = [0, 1]$. Thus $\mu(E) = 1$. As in Equation 2.14 from Example 2.8, define the contractions of the iterated function system as

$$S_1 = \frac{x}{3} \text{ and } S_2 = \frac{x}{3} + \frac{2}{3}.$$

Let $k \in \mathbb{N}$, and consider the approximation of the Cantor set, F_k. We approximate a mass distribution on the Cantor set by studying the intermediate mass distribution on F_k,

as in the example in Section 2.3 and as detailed in Falconer [36]. At the kth stage in the construction of the Cantor set the subintervals consists of the union of all k-length compositions of the contractions S_1 and S_2. The information about a particular subinterval from this collection is recorded in the k-length vector $(i_i, \ldots, i_k)$, where $i_j \in \{1, 2\}$. Each subinterval is of the form $S_{i_1 \cdots i_k}(E) := S_{i_1} \circ \cdots \circ S_{i_k}(E)$, where $i_1 \cdots i_k$ is shorthand for $(i_i, \ldots, i_k)$. Hence, we write F_k as the union of all of the

subintervals,

$$F_k = \bigcup_{I_k} S_{i_1 \cdots i_k}(E),$$

where $\mathcal{I}_k = \{i_i \dots i_k \,|\, i_j = 1, 2\}$.

As described in Section 2.3, we use the concept of a mass distribution in order to introduce a multifractal measure on the Cantor set. Begin by choosing $p_1, p_2 > 0$ such that $p_1 + p_2 = 1$. For a fixed p_1 and p_2 the mass distribution for the first and second levels are

$$\mu(S_1(E)) = p_1 \qquad \mu(S_2(E)) = p_2$$

$$\mu(S_1^2(E)) = p_1^2 \quad \mu(S_{12}(E)) = p_1 p_2 \qquad \mu(S_{21}(E)) = p_2 p_1 \quad \mu(S_2^2(E)) = p_2^2.$$

Let the number of subintervals be $L_k = 2^k$. Enumerate the subintervals at the kth level so that $F_k = \bigcup_{i=0}^{L=2^k} E_i$. Each subinterval E_i can be covered by a ball of diameter 3^{-k} and has a mass of

$$\mu(E_i) = p_1^k p_2^{k-l}, \tag{4.1}$$

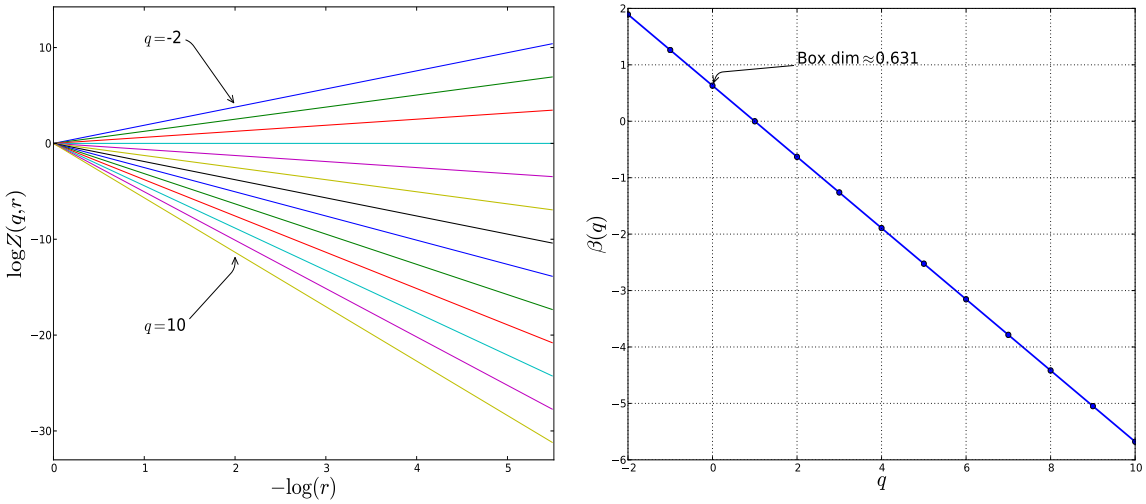
where l denotes the number of 1's in $(i_1, i_2, \dots, i_k)$.

4.1.2 Analysis of the Cantor Set

From the above framework we are able to analytically compute the multifractal spectrum using the partition function methods described in Subsection 2.3.5.1 using the measures of the subintervals from Equation 4.1. Recall the definition of the partition function for a fixed resolution, or radius, r and the parameter $q \in \mathbb{R}$:

$$Z(q, r) = \sum_{B\,:\,\mu(B)>0} \mu(B)^q.$$

In Figures 4.1(a) and 4.2(a), the numerous partition functions are plotted for values the parameter q. Both figures show $Z(q, r)$ for $q \in \{-2, -1, \dots, 10\}$ in order from



(a) Partition Functions

(b) $\beta(q)$ Curve

Figure 4.1: (a) Partition functions determined analytically for $p = 0.5$ and values of $q \in \{-2, -1, \ldots, 10\}$. The rate of change of the slopes between successive q values is linear, as shown in Lemma 4.1. This observation implies that the $\beta(q)$ curve is linear with a slope of $\frac{\log 2}{\log 3}$.

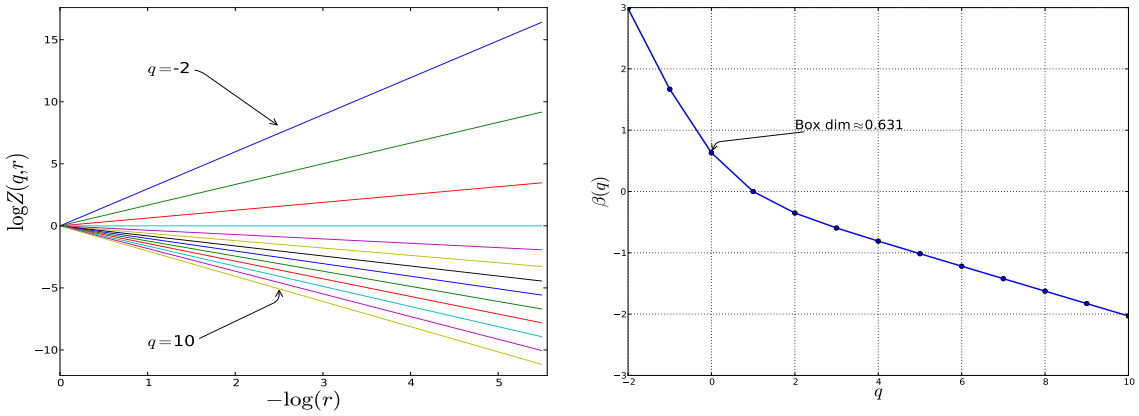
top to bottom. Recall that the slope of the linear regression curve to each partition function yields a single value, $\beta(q)$, which is plotted against q to obtain Figures 4.1(b) and 4.2(b).

Note that when $q = 0$,

$$Z(0, r) = \sum_{B\,:\,\mu(B)>0} \mu(B)^0 = \#\{B \,:\, \mu(B) > 0\}, \tag{4.2}$$

which simply counts the number of balls necessary to cover F_k. Thus, as previously noted in Section 2.3, when the measure is excluded from the calculation and we obtain the box-counting dimension of the Cantor set, which in both figures is $\frac{\log(2)}{\log(3)}$. This can be seen in both Figures 4.1(b) and 4.2(b) where $\beta(0) = \frac{\log(2)}{\log(3)}$

We are interested in the accuracy of our algorithms. We begin with the Cantor set with $p_1 = p_2 = 0.5$. In this case we obtain the uniform Cantor set with no complicated distribution of mass since $\mu(S_{i_1 \cdots i_k 1}(E)) = \mu(S_{i_1 \cdots i_k 2}(E))$ As can be seen



(a) Partition Functions

(b) $\beta(q)$ Curve

Figure 4.2: (a) Spectrum of partition function curves determined analytically for $p_1 = 0.2$ and values of $q \in \{-2, -1, \ldots, 10\}$. (b) The $\beta(q)$ curve. For each q, the slope of the linear regression curve to $Z(q, r)$ is computed to give $\beta(q)$. Note that for $q = 0$ we get the box counting dimension of the set, $\log(2)/\log(3) \approx 0.631$.

in Figure 4.1(a), the slopes of the graphs of the $-\log(r)$ vs. $\log(Z(q, r))$ change linearly with q. This is a consequence of the following lemma.

Proposition 4.1. *Let F be a Cantor set constructed with the contraction S_i, where $i = 1, 2, \ldots, m$ and the contraction ratios are $c_i = c^{-1}$ for $c > 1$. Suppose that the mass is distributed evenly such that $\mu(S_i(I)) = \dfrac{\mu(I)}{m}$ for each subinterval I. Then for $q \in \mathbb{R}$ the slope of $Z(q, r)$ can be determined by*

$$\beta(q) = (1 - q)\frac{\log m}{\log c}.$$

Proof. Consider the kth level in the construction of F and let $r_k = c^k$ be the radius necessary to cover each subinterval by a ball B_{r_k}. Then $\mu(B_{r_k}) = \frac{1}{m^k}$ since the mass is distributed even amongst each subinterval at each level. Let $x = \frac{1}{m}$ then the rate

of change between r_k and r_{k+1} is

$$\begin{aligned}\frac{\log\left(\sum_{i=0}^{m^{k+1}} x^{(k+1)q}\right) - \log\left(\sum_{i=0}^{m^k} x^{kq}\right)}{-\log(c^{-(k+1)}) + \log(c^{-k})} &= \frac{1}{\log c}\log\left(\frac{m^{k+1}(x^{kq+q})}{m^k(x^{kq})}\right) \\ &= \frac{\log(mx^q)}{\log c} \\ &= \frac{\log(m^{1-q})}{\log c} \\ &= (1-q)\frac{\log m}{\log c}.\end{aligned} \tag{4.3}$$

□

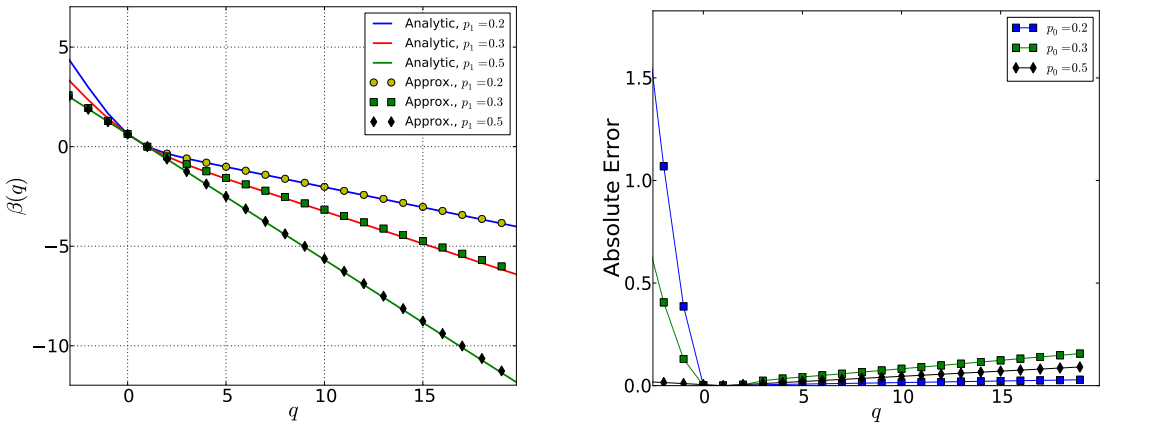
For the traditional, monofractal ternary Cantor set, $m = 2$ and $c = 3^{-1}$. Thus, using the linear regression slope at each q to approximate $\beta(q)$ we obtain,

$$\beta(q) = (1-q)\frac{\log 2}{\log 3} = \frac{\log 2}{\log 3} - q\frac{\log 2}{\log 3}. \tag{4.4}$$

This is exactly the line $\beta(q)$ in Figure 4.1(b). This is indicative of the single scaling exponent of $^{\log 2}/_{\log 3}$. In contrast to this single scaling exponent, when $p_1 = 0.2$ and $p_2 = 0.8$ we obtain nonlinear changes between the slopes of the partition functions, which yields the nonlinear $\beta(q)$ curve in Figure 4.2(b).

Different values of p_1 and p_2 yield the "spectrum" of multifractal spectrum curves in Figure 4.3(a). As $p_1 \to 0.5$, the corresponding β curves approach the linear curve of the monofractal Cantor set. For values of q between 0 and 10, the fit of the approximate $\beta(q)$ curve to the analytically determined curve is excellent, as can be seen in Figure 4.3. Yet, the approximation of the multifractal spectrum loses its tight fit for $q < 0$. This is especially clear in Figure 4.3(b), and is a common problem encountered in multifractal research. The issue is simple: When $q < 0$, the regions with the least measure have a tendency to diverge [36]. To our knowledge, the method that has proven most successful in accurately computing the multifractal spectrum

for $q < 0$ is the wavelet transform modulus maxima method (see Section 2.4 or [2]). Nevertheless, $q < 0$ falls outside of the the thermodynamic analogy, since physically this implies a negative temperature.



(a) Analytic vs. Approximate Spectra

(b) Difference between Analytic and Approximate Curves

Figure 4.3: (a) Analytic (solid lines) and approximate (markers) multifractal spectra for Cantor sets with different mass distributions. (b) Curve correspond to error between analytic and approximate curve for each p_0 in (a). For $0 \leq q \leq 10$ the approximate curves are highly accurate compared to the analytic curves. For $q < 0$, the error increases. Past research has encountered similar problems [6, 43].

It is useful for the sake of computational analysis to plot the $\beta(q)$ curve for the analytic and approximate cases. Yet, the ultimate goal of a multifractal analysis is to understand the dimensions of the interwoven sets

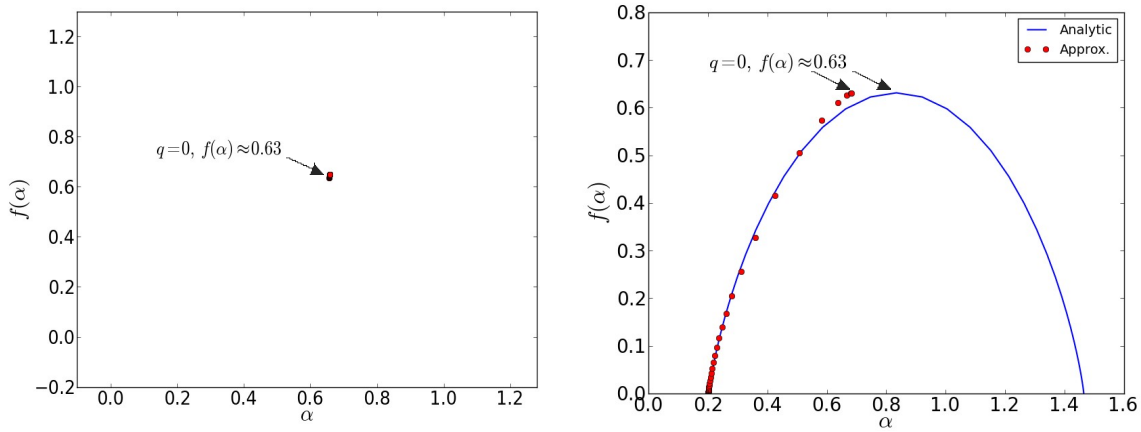
$$F_\alpha = \{x : \dim_{\text{loc}} \mu(x) = \alpha\}, \tag{4.5}$$

over a range of α's. Recall (Section 2.3.1) that

$$f(\alpha) := \dim_H F_\alpha.$$

We apply the simplicial measure to Equations (2.27) and (2.28) in order to compute the $f(\alpha)$ spectrum for both mono- and multifractal measures on Cantor sets. For

instance, the ternary Cantor set possesses a single scaling exponent, namely $\alpha^* = \frac{\log 2}{\log 3}$. This succinctly describes the fact that the measure is concentrated entirely on the set $F_{\alpha^*} = \{x : \dim \mu(x) = \alpha^*\}$. In addition, the dimension of F_{α^*} is $f(\alpha^*) = \frac{\log 2}{\log 3}$. Analytically this yields an $f(\alpha)$ spectrum that is a single point at (α^*, α^*). Given an approximation of the ternary Cantor set, numerical inaccuracy due to the finite set of points causes the graph of the $f(\alpha)$ spectrum in Figure 4.4(a) to spread a small amount. Recall that (2.27) and (2.28) are implicit one-parameter functions of q. When $q = 0$ we are simulating the box-counting method (Section 2.3.2 on page 14), and expect that $(\alpha(0), f(0)) \approx 0.631$. This is indeed the case, as pointed out in Figure 4.4(a). To obtain this plot we computed f and α for values of q between 0 and 10.



(a) $p_0 = 0.5$ (b) $p_0 = 0.2$

Figure 4.4: Approximation of the $f(\alpha)$ curve using Equations (2.27) and (2.28). (a) A Cantor set with approximately 8000 points with $p_0 = 0.5$. The fractal dimension for $q = 0$ is approximately 0.631. (b) The approximation (red circles) of the $f(\alpha)$ curve for a multifractal measure on the Cantor set, constructed with $p_0 = 0.2$ and approximately 12,000 points. The analytically-computed curve (blue line) is plotted for comparison ($-10 < q < 10$). The values for $q = 0$ are marked. Both curves are computed for positive values of q between 0 and 10.

We next confirm that the simplicial measures detect the wider range of scaling exponents present in a Cantor set supporting a multifractal measure. With $p_0 = 1 - p_1 = 0.2$ we obtain the $f(\alpha)$ spectrum (red circles) in Figure 4.4(b) for positive q. (The approximate curve uses the same range for q as above.) We overlay this onto this the $f(\alpha)$ curve plotted by analytically calculating the mass of each interval. For this curve, value of negative q are stable and so we can fill out the other half of the curve. We conclude from the results in this section that the simplicial measure accurately captures the multifractal properties in the case of measures concentrated on self-similar Cantor sets.

Notice that the dimension estimation is quite accurate. We find a small amount of error in the values for α, especially when q is near zero (see Figure 4.4(b)). Recall that the implicit value for α is actually obtain by computing the expected value $\langle\alpha\rangle$. This is a issue currently under investigation. We will encounter this again when we extend these results to the more complex case of the Hénon attractor.

4.2 The Henon Map

The Henon map is a nonlinear map in two dimensions whose simple definition belies its complexity. It is defined by

$$\begin{aligned} x_{n+1} &= 1 - ax_n^2 + y_n \\ y_{n+1} &= bx_n, \end{aligned} \tag{4.6}$$

for $a, b \in \mathbb{R}$. Since its inception (see M. Hénon in [46]), it has provided an often-used test bed example for both theoretical and computational multifractal techniques. For certain parameter values the map is chaotic, including the canonical values $a = 1.4$ and $b = 0.3$ which are the most studied [6, 43]. Indeed, the parameter space itself

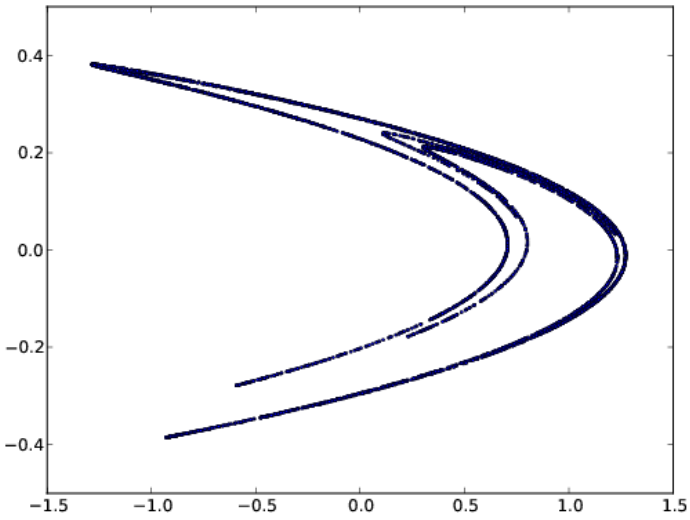
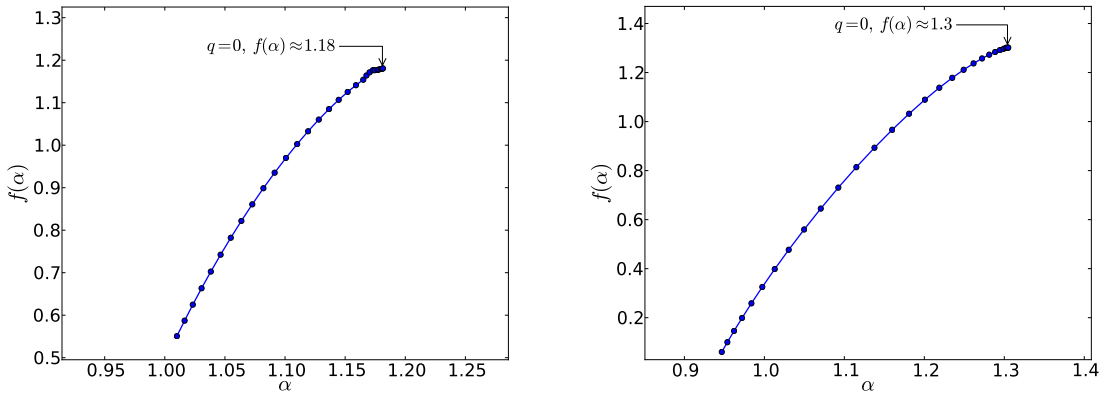


Figure 4.5: Henon map with 5000 iterations with $a = 1.4$, $b = 0.3$.

yields an incredibly rich structure and has been a fruitful area of study [12, 14, 72]. Despite years of research, there is no analytic calculation of the multifractal spectrum of the natural measure on the Hénon attractor.

We approximate the Hénon attractor by generating N iterates of (4.6). Though we have effectively applied the simplicial measure and tidy cover algorithms to data sets containing upwards of 50,000 points, we show results for $N = 5000$ and $N = 8000$. The ranges of scales investigated are $2^{-10} \leq \varepsilon \leq 2^{-3}$ and $2^{-10} \leq \varepsilon \leq 2^{-4}$, respectively. These approximations of the Hénon attractor consist of a relatively small number of points. In fact, our results are impressive when compared to Arneodo's [6], in which one billion iterates were used. Arneodo, *et al.*, obtain estimates of $f(\alpha(0)) \approx 1.26$, whereas the Kaplan-Yorke estimate of the dimension of the Hénon attractor gives $f(\alpha(0)) = 1.276 \pm 0.001$ [11]. Using the implicit technique described in Section 2.3.5.2, we compute an estimate of the $f(\alpha)$ spectra obtained using simplicial measures on 5000 and 8000 points and plot the result in Figure 4.6. We considered resolutions

$2^{-10} \leq \varepsilon \leq 2^{-4}$. For the case with 5000 points we clearly do not have the required accuracy. Increasing the number of iterates to 8000 brings the estimate of the "capacity dimension", when $q = 0$, to approximately 1.3, which is within 0.03 of the Kaplan-Yorke estimate. These results agree well with past results [6, 43], while using up to five order of magnitude fewer points. Of equal importance, the results show the efficacy of using the simplicial measure to study the SBR measure on attractors with scaling properties beyond those of "cookie cutter" sets such as the Cantor set.



(a) $N = 5000$ (b) $N = 8000$

Figure 4.6: The $f(\alpha)$ singularity spectrum for an approximation of the SBR measure on the Henon attractor.

An important issue to note is the range of smaller range of scaling exponents α in the $f(\alpha)$ curve. In Figure 4.7 we have overlayed the $f(\alpha)$ curve from Figure 4.6(b) (blue circles) with an image of the $f(\alpha)$ curve plotted in Arneodo [6] (black line). (The two outer lines in the Arneodo plot are from various approximations. See [6] for more details.) At the apex of the curve, we see that the dimension of the set for which $\alpha(0) \approx 1.3$ is similar to that found by Arneodo, *et al.*. The range of α computed using the simplicial measure on 8000 points, though, is not as great (does not extend as far

to the left) as the range for α in Arneodo's work. Nevertheless, the approximation near the apex of the curve, for positive values of q near 0, is impressive.

4.3 Neural Networks

Much of the work presented in the previous chapters resulted from a desire to analyze a certain mathematical model of a network of neurons and understand its "robustness". We describe the mathematical model as a proxy for a physiological system below, as well as the motivations for studying it. The notion of a robust versus a compromised system is inherent to this study. We touch on this first, then describe various techniques that our collaborators have developed to study the robustness of a physiological systems. These tools have proven effective at classifying real-world dynamical systems. Nevertheless, they proved less than ideal for optimization of a mathematical model. This led to exploration of new techniques, culminating in the simplicial measure technique in Chapter 3. In the following subsections we describe results obtained for mathematical models of neural networks, and discuss some of the difficulties encountered when trying to extend these.

4.3.1 Motivation For studying Neural Networks

Measuring an output of a physiological system provides a window into its complex multi-scale dynamics. The measurements are often spread over time and many techniques of time series analysis have been used to gain an insight into the underlying physiology. Some of the most intriguing observations indicate that the complexity of the time series produced by a free-running physiological system, as measured by Detrended Fluctuation Analysis (DFA), Multiscale Entropy (MSE) and other methods, is correlated to the robustness and health of the physiological system. More precisely,

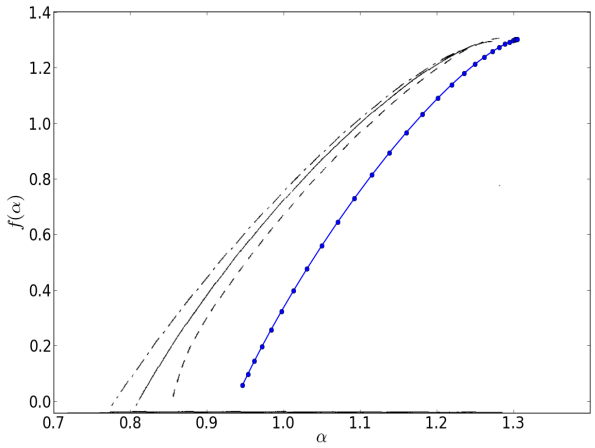


Figure 4.7: Overlay of the $f(\alpha)$ spectrum from Figure 4.6(b) (blue circles) and the $f(\alpha)$ spectrum from Arneodo, *et al.*, "Fractal dimensions and $f(\alpha)$ spectrum of the Henon attractor" [6] (solid black line). This is the result of 10^8 iterates of the Hénon map and was obtained by computing the Legendre transform of the D_q spectrum (see Section 2.3.5.1).

analysis of time series gathered from the measurement of cardiac inter-beat intervals, oscillations of red blood cells, gait analysis, and other patterns observed in living organisms suggests that healthy systems produce complex time series, while compromised systems produce either very simple periodic signals, or completely random signals [26, 28, 41, 40, 66, 64].

The potential diagnostic and therapeutic consequences of this hypothesis demand studies that go beyond passive analysis of existing data. What is needed is a model which reproduces observed characteristics of physiological signals and thence can be actively tested. Ideally, such a model would be based on our knowledge of a complex system. However, it has proved challenging to construct a dynamical system-based model that reproduces the statistical characteristics of physiological time series. The only successful attempt was a discrete map with added noise which partially reproduced some of these characteristics [65].

The central aim of our work is to construct a deterministic FitzHugh-Nagumo-based neural network model which exhibits the complex signatures measured by DFA and MSE metrics in physiological signals. We study the dependence of these metrics on the length of the computed time series and initial conditions used. Note that this issue has relevance to the analysis of experimental time series. One tacitly assumes that the analyzed series represents the steady response of the system, which does not depend on initial data, or the time when the measurements started. The longer the time series, the more time the system has for the initial data effect to "average out". While these assumptions are satisfied if the deterministic system is ergodic and stationary, we can test them directly in our model.

The model consists of a network of five excitatory cells and five inhibitory cells. The structure and dynamics of the network are based on the work of Terman [74] modeling the structure of the hippocampus. Analyzing the time series of averages from the excitatory cells' voltage potential, we show that it matches the DFA and MSE measurements of complexity for free-running physiological systems across a range of time scales. These results do not depend on the initial conditions used. Furthermore, we show that the range of complex behavior grows when we increase the length of the time series from $15,000$ to $100,000$ (in arbitrary units), but does not grow further, when we extend the length to $400,000$.

This indicates that the system has a certain capacity for complexity which does not depend on initial conditions, and which is recovered from data of finite length.

4.3.2 Neural Network Model

We create a bipartite graph $G = \langle V_E, V_I, A \rangle$, where V_E consists of excitatory nodes, or cells, V_I consists of inhibitory cells, and A is the set of directed edges which

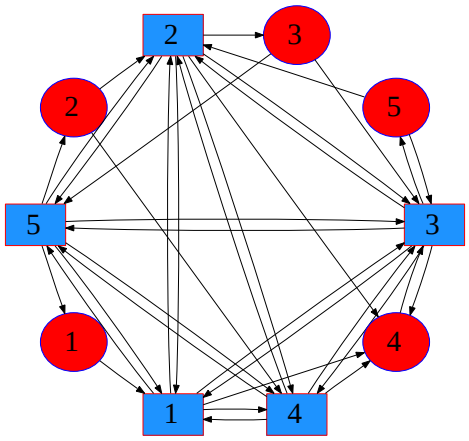


Figure 4.8: A neural network with ten cells. Inhibitory cells (I) are represented by blue squares, excitatory cells (E) by red ellipses. The subgraph of I cells is completely connected. $E \to I$ and $I \to E$ edges are created with probability $\frac{\ln N}{N}$. $E \to E$ edges are not allowed.

consists of the following types of connections:

$$e_j \to i_k, \quad i_k \to e_j, \text{ and } i_k \to i_{k'}, \ k \neq k',$$

where e denotes an excitatory cell and i denotes an inhibitory cell; $j \in \{1, \ldots, n\}$ and $k, k' \in \{1, \ldots, m\}$,and $n = |V_E|$, $m = |V_I|$. Excitatory cells do not connect to other excitatory cells in order to avoid the blow-up of the solutions due to runaway positive feedback. The subgraph V_I is complete. We construct the remaining connections randomly by selecting the $e_j \to i_k$ and $i_k \to e_j$ edges with probability $\rho = (\ln N)/N$, where $N = n + m$. Figure 4.8 shows a sample neural network. Edges are weighted by a maximal conductance constant which depends on the type of connection. We represent these weights by g_{IE}, g_{EI}, $g_{II} \in (0, 1)$, where the subscripts E and I denote excitatory and inhibitory edges, respectively, and g_{xy} specifies the weight for the directed edge $x \to y$.

4.3.3 FitzHugh-Nagumo Equations

The following system of coupled differential equations describes the behavior of each cell in the graph defined above [74]. All units are arbitrary.

Inhibitory cells:

$$\begin{aligned}
\frac{dv}{dt} &= v - \frac{v^3}{3} - w - g_{II}(v - v_I)\left(\sum s_k\right) \\
&\quad - g_{EI}(v - v_E)\left(\sum s_j\right) + K_I(t) \\
\frac{dw}{dt} &= \epsilon(v - bw + c) \\
\frac{ds}{dt} &= \alpha_I(1 - s)h(x, \theta_x) - \beta_I s \\
\frac{dx}{dt} &= \epsilon[\alpha_x(1 - x)h(v, \theta_I) - \beta_x\, x]. \qquad (4.7)
\end{aligned}$$

Excitatory cells:

$$\begin{aligned}
\frac{dv}{dt} &= v - \frac{v^3}{3} - w - g_{IE}(v - v_I)\left(\sum s_k\right) + K_E(t) \qquad (4.8) \\
\frac{dw}{dt} &= \epsilon(v - bw + c) \\
\frac{ds}{dt} &= \alpha(1 - s)h(v, \theta) - \beta s, \qquad (4.9)
\end{aligned}$$

In general, the coupling variable s represents the fraction of open synaptic channels. The coupling sums, $\sum s_k$ and $\sum s_j$, are limited to those cells connecting to the given cell; s_k is the input from an inhibitory cell, s_j is the input from an excitatory cell.

A direct synapse is one in which the postsynaptic receivers are themselves ion channels. In equations (4.9) an excitatory cell is modeled with a direct synapse. The function h is a steep sigmoidal curve allowing for a very rapid, but still continuous, activation of the synaptic processes. Once the voltage potential v crosses θ the synapse activates (h turns on) and an impulse travels to connected cells.

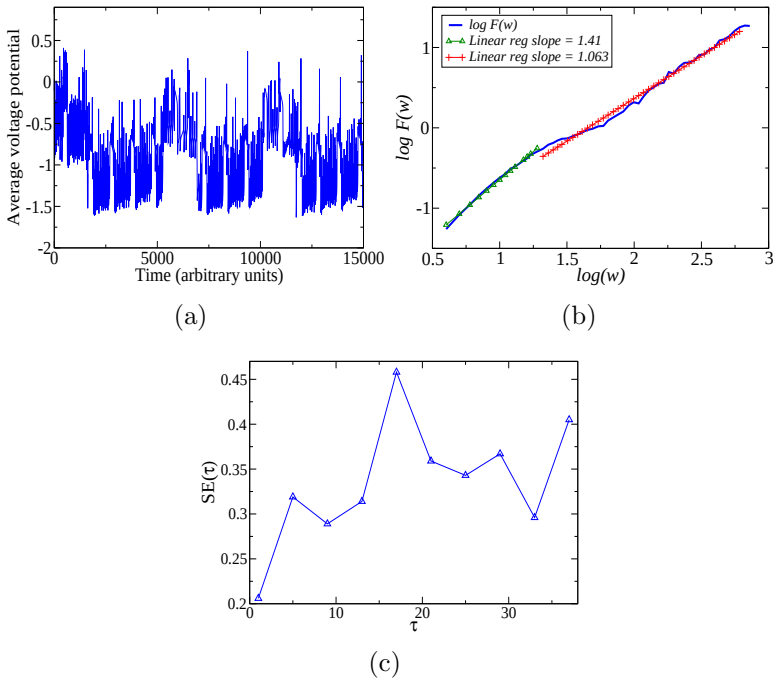


Figure 4.9: (a) Time series for a simulation run for 15000 "seconds" (units are arbitrary). (b) The DFA curve exhibits two trends: at short scales we observe autocorrelation effects due to deterministic ODE solver, at longer scales we find a scaling exponent of $\beta = 1.063$, similar to physiological systems, though over a shorter range. The relatively short time series produces more variability in the higher window sizes since the number of data points used to calculate the DFA becomes small. (c) Sample entropy $SE(\tau) = SE(m, \delta, N, \tau)$ is calculated for $\tau = 1, 5, \ldots, 37$, $m = 7$, and $\delta = 0.2$. The length, N, of the coarse-grained time series depends on τ. The MSE curve has an average of $\overline{SE} = 0.35$. The variance of $Var(SE) = 0.003$ is larger than in longer simulations. As in (b), this greater variability is due to the relatively small number of data points in the original time series.

An indirect synapse, where the postsynaptic receivers are not ion channels, is modeled by adding the delay variable x. All inhibitory synapses in our model are indirect. The activation of the synaptic variable s relies on x, not v as in a direct synapse. In the bottom equation of (4.7), v must first reach the threshold θ_I in order to activate x. After this delay, x goes on to activate s. Each cell is assigned a unique $\varepsilon \ll 1$. If the cell oscillates when disconnected from the other cells, ε is inversely proportional to the period of this oscillation. The constants v_E and v_I are reversal

potentials for excitatory and inhibitory cells, respectively. The maximal conductance constants $g_{IE},\ g_{EI},\ g_{II}$ weight a cell's input by multiplying the coupling terms.

The parameters $\alpha, \alpha_I,\ \alpha_x,\ \beta,\ \beta_I,\ \beta_x \in \mathbb{R}$ are rates at which the synaptic variables, s and x, turn on and off. The linear recovery term is specified by the parameters $b,\ c \in \mathbb{R}$. And the K_E and K_I terms represent external input to the system. In order for cells to enter an excitatory state a small amount of input must be applied to the system. For simplicity we use constant input. The complete set of parameters used is listed in the Appendix.

4.3.4 *In-Silico* Neural Network

An *in-silico* neural network, $\mathcal{N}$ is constructed from the following constituent pieces. The structure and behavior of the network is specified in a graph G composed of cells whose behavior is described by (4.7) and (4.9). A parameter set P contains fixed parameters for the FitzHugh-Nagumo equations. Thus, $\langle G, P \rangle$ generate a unique neural network $\mathcal{N}$ used for simulations.

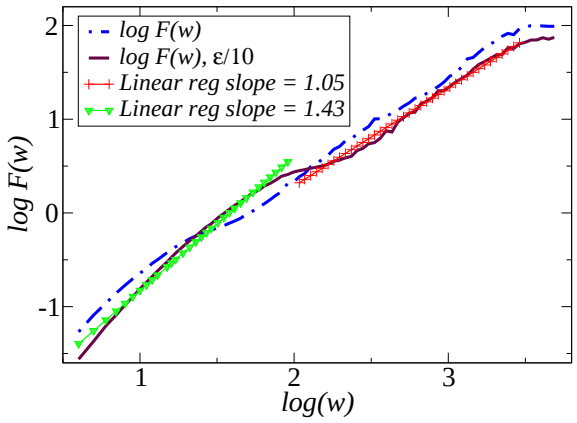


Figure 4.10: DFA curve from an initial ε distribution (solid blue), and the DFA curve obtained by dividing each ε by 10 (dashed purple).

4.3.5 Time Series Analysis

Solving the above system of ODE's with an adaptive step solver results in a set of solutions for each cell. The voltage potentials, $v_j(t)$, from all excitatory cells are averaged at each time step to give a time series $g(t)$. The analysis techniques require equally spaced time steps. Since an adaptive step algorithm returns irregularly spaced time steps we construct a new time series by partitioning the time axis into *bins* and averaging over these bins, giving a time series composed of the values

$$\bar{g}(n) = \frac{1}{\kappa T} \sum_{i=n(\kappa T)}^{(n+1)(\kappa T)} g(i), \tag{4.10}$$

where $n = 0, 1, \ldots, L = \lfloor \frac{M}{\kappa T} \rfloor$, M is the length of $g(t)$, and κ is chosen so that each bin contains a minimum number of points. Note that a small tail of the original time series must be discarded. We apply the following techniques to series $\bar{g}(n)$.

4.3.6 Detrended Fluctuation Analysis

Detrended fluctuation analysis is a statistical method developed to determine long term trends in time series [63, 64]. Given a time series of length N, it is first integrated at each point to give a function $y(t) = \int_0^t g(s)\, ds$. The time axis is then partitioned into windows of length w. Next, a linear regression line, $y_w(t)$, is fit to the integrated curve for each window. The root-mean-square of the *detrended* curve $y(t) - y_w(t)$ is calculated, giving the detrended fluctuation value for a window size w:

$$F(w) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left(y(k) - y_w(k)\right)^2},$$

where y_w is understood to be the linear regression to y defined piecewise over each window of length w. F is computed for a wide range of window sizes and typically

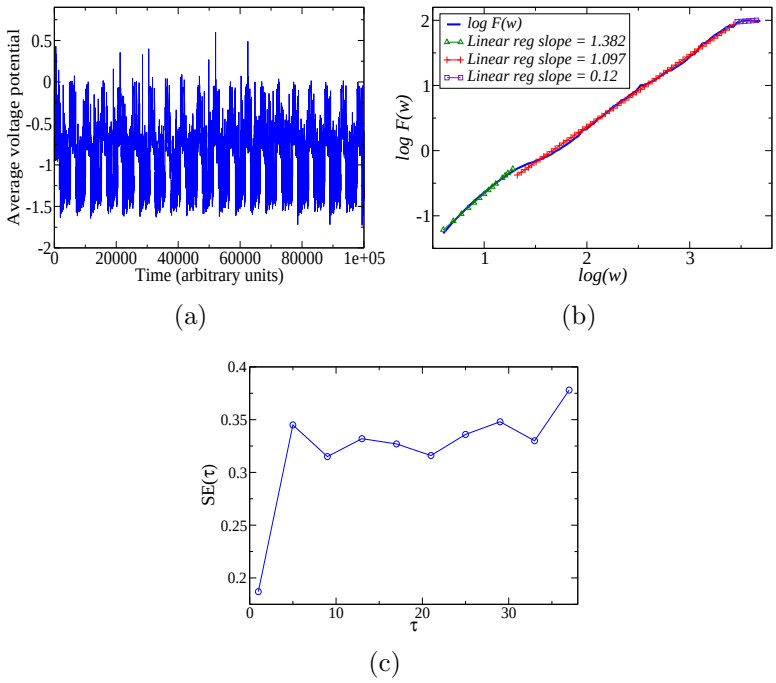


Figure 4.11: (a) Time series from simulating $\mathcal{N}$ for 100000. (b) The region of physiological complexity for the DFA curve extends from $w = 10^{1.3}$ to $w = 10^{3.5}$. For windows larger than $10^{3.5}$ we observe a new, flat trend indicating that the time series generated by $\mathcal{N}$ is devoid of long range correlations for these larger windows. (c) The MSE curve is relatively constant over a large number of scales indicating physiological complexity in the time series. The average sample entropy for $\tau \geq 5$ is $\overline{SE} = 0.34$, with a variance of $Var(SE) = 0.0007$.

increases monotonically. Power law scaling exists in the time series when a log-log plot produces a linear relationship.

4.3.7 Multiscale Entropy

Multiscale entropy [27, 28] simulates the sequence of refinements in the definition of Kolmogorov-Sinai (KS) entropy [52]. In the case of MSE, though, we are interested in the evolution of the entropy across these refinements, and not their limit. Suppose we obtain a time series $g(t)$ by taking measurements of an experiment. This gives a sequence of data points $\{g(0), g(1), \ldots, g(N)\}$. MSE simulates the situation where

we perform the identical experiment with less time accuracy. A time series for this situation is constructed through *coarse graining*, or partitioning the time axis of the original series into blocks of size $\tau \in \mathbb{N}$ and averaging the data over these windows. Thus, each coarse grained time series is composed of the points

$$g_\tau(n) = \frac{1}{\tau} \sum_{k=n\tau}^{(n+1)\tau} g(k),$$

where $n = 0, 1, \ldots, L = \lfloor N/\tau \rfloor$.

The entropy of this new time series $\{g_\tau(0), g_\tau(1), \ldots, g_\tau(L)\}$ is estimated using *sample entropy* [68]. Sample entropy views a time series as a sequence of random variables and measures the creation of information by computing the correlation between delay vectors of length m and $m+1$.

In order to define sample entropy, fix τ and set $g_\tau(i) = x_i$. Given m, let $u_m(i) = \{x_i, x_{i+1}, \ldots, x_{i+m}\}$ be a delay vector of length m, and define the number of vectors close to $u_m(i)$ as $n_i^m(\delta) = \#\{x_m(j) : d(x_m(i), x_m(j)) < \delta\}$ where $\delta > 0$ is some tolerance and

$$d(u_m(i), u_m(j)) = \max_{0 \le k < m} \{ \mid x_m(i+k) - x_m(j+k) \mid \}.$$

There are $N(m) = L - m$ full length vectors $u_m(j)$, excluding the possibility of self matches. The probability of finding the vector $u_m(j)$ within a distance δ of $u_m(i)$ is

$$C_i^m(\delta) = \frac{n_i^m(\delta)}{N(m)}. \tag{4.11}$$

For the parameter m the probability that any two vectors are within δ of each other is

$$C^m(\delta) = \frac{1}{N(m)} \sum_{i=0}^{N(m)} C_i^m(\delta) \tag{4.12}$$

The above *correlation integral* is used to define the sample entropy for the delay m, tolerance δ, and time series length L as

$$SE(m, \delta, L) = -\ln \frac{C^{m+1}}{C^m}.$$

C^{m+1}/C^m is commonly thought of as the information gained as the trajectory moves from time $m\tau$ to $(m+1)\tau$. A larger difference between C^m and C^{m+1} results in more information, i.e., a higher value of SE. For a fixed value of m the graph of the sample entropy over a range of τ's provides a measure of the amount of long range correlation in the time series. A relatively constant amount of entropy across many values of τ signifies correlations amongst data points over multiple time scales. For instance, $1/f$ noise which is highly correlated across time scales yields a constant MSE curve. In contrast, white noise is monotonically decreasing since it possesses no long range correlations.

4.3.8 Results

Consider an *in-silico* neural network $\mathcal{N}$ as defined in Section 4.3.4. We construct the time series of voltage potentials $\bar{g}(n)$ by choosing a set of initial conditions and solving the system of differential equations over the time interval $[0,\ N]$, then performing the binning procedure described in (4.10). The units of time are arbitrary.

Figure 4.9 shows the time series, DFA curve, and MSE curve for a simulation of $\mathcal{N}$ run for $N = 15000$. In Figure 4.9(b), the scaling exponent $\beta = 1.063$ over the range of window sizes $w = 10^{1.3}$ to $w = 10^{2.7}$. Thus, running $\mathcal{N}$ for a relatively short simulation produces power law scaling similar to a physiological system over the range $(1.3, 2.7)$ for the total length of range 1.4. This range of scales where $\beta \approx 1$ is shorter than that typically seen in biological systems, where the range typically has length greater than 3. The DFA curve extracted from our simulations has three distinct

regions. In the first region, where $w < 10^{1.3}$, the linear regression deviates from the power law $\beta \approx 1$ due to autocorrelation effects on short scales, which are caused by the deterministic ODE solver. These effects dominate at scales much smaller than the highest frequency cellular oscillations, which can be estimated from the largest ε. Figure 4.10 illustrates this effect; we compare the DFA curve in Figure 4.11(b) to the curve obtained after dividing by 10 all of the ε's used in generating Figure 4.11(b). The deterministic portion extends to higher scales because the highest oscillation frequency decreased by a factor of ten. This also illustrates the importance of the choice of ε's on the DFA curve. In order to avoid these deterministic effects we focus (for the original set of ε's listed in the Appendix) on power law scaling in the second region, where $w > 10^{1.3}$.

In Figure 4.9(b) we see that $\mathcal{N}$ produces physiologically complex behavior in the region $w > 10^{1.3}$, which ends at $w = 10^{2.7}$. By increasing the length of the time series (Figure 4.11(b)) the second region extends past $w = 10^{2.7}$ to $w = 10^{3.5}$ showing that $\mathcal{N}$ continues to introduce complexity into the time series past the scale limits imposed by the short simulation in Figure 4.9. The second region terminates in Figure 4.11(b)

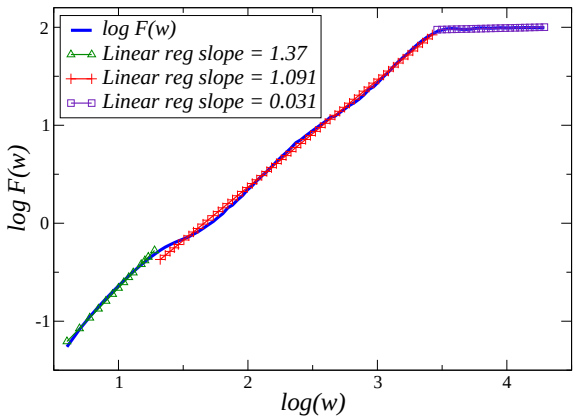


Figure 4.12: The DFA curve after simulating $\mathcal{N}$ for $N = 400000$. The region of complexity remains unchanged from that seen in Figure 4.11(b). This implies that $\mathcal{N}$ has an inherent limit for generating long range correlations.

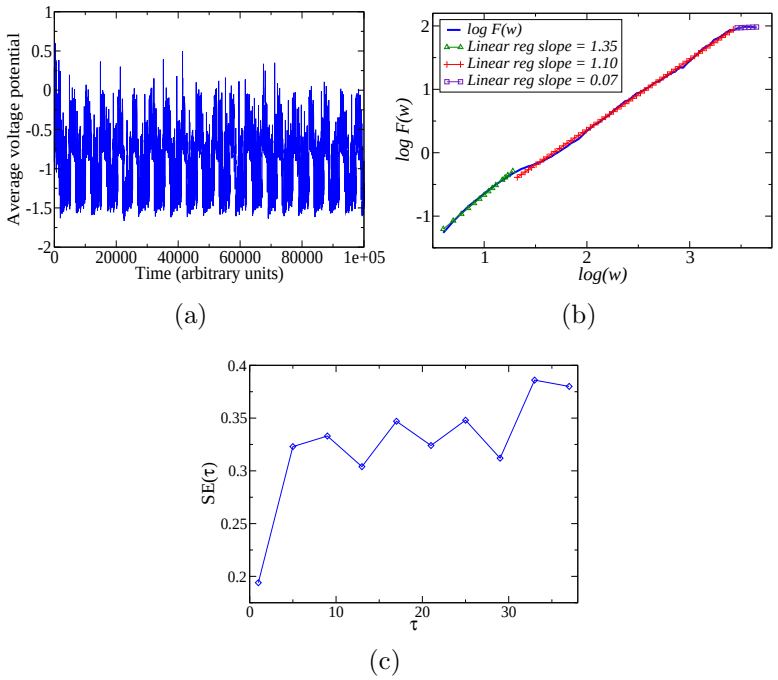


Figure 4.13: (a) Time series resulting from choosing a different initial condition for the excitatory cell e_1 compared to Figure 4.11 and simulating $\mathcal{N}$ for $N = 100000$. (b) The DFA curve exhibits a nearly identical scaling exponent $\beta = 1.10$ over the middle region as the curve in Figure 4.11(b). Long range correlations are unaffected by initial conditions. (c) The MSE curve has slightly different values at the various scales, but the average entropy for $\tau \geq 5$ is $\overline{SE} = 0.34$ which is identical to that produced by $\mathcal{N}$ in Figure 4.11. The variance $Var(SE) = 0.0003$ is similar as well.

at $w = 10^{3.5}$, where a third region with no long range correlation begins. Extending the length of the simulation to $N = 400000$ yields the DFA curve in Figure 4.12, where this third region extends to larger window sizes. Clearly the extension of the time series fails to find longer range correlations in the time series. We conclude that the system $\mathcal{N}$ has an upper limit $w = 10^{3.5}$ on the length of long range correlations it can generate.

The longer time series yields an MSE curve that is relatively constant, mimicking the behavior observed in simulation of $1/f$-noise as well as free-running physiological systems. It maintains an entropy level nearly identical to the MSE curve in Fig-

ure 4.9(c). Indeed, the average entropies for $\tau \geq 5$ are $\overline{SE} = 0.34$ and $\overline{SE} = 0.35$, respectively. region where the MSE curve is most nearly constant with the least amount of variation.) The MSE curve in Figure 4.9(c) derived from the shorter time series suffers larger variations due to coarse graining effects on the relatively low number of data points in the original series. Nevertheless, as the comparison of the averages shows the MSE and sample entropy measures for shorter simulations are consistent with the results from longer simulations, and still provide good insight into the complexity of the network.

Furthermore, the behavior of $\mathcal{N}$ does not depend on initial conditions as we confirm by choosing random initial conditions for excitatory cells uniformly in the interval $(-5, 5)$. To illustrate, we present a typical case in which we alter the initial condition of one excitatory cell, e_1. Figure 4.13 shows the time series and related analysis obtained by using a random initial condition. Comparison of the time series in Figure 4.13(a) to the one in Figure 4.11(a) shows minor differences. Importantly, the scaling exponent for the DFA curve in Figure 4.13(b) differs from that in Figure 4.11(b) by less than 0.01. The mean value of the entropy for both MSE curves is $\overline{SE} = 0.34$. Figure 4.14 shows a comparison of the MSE curves from Figures 4.11(c) and 4.13(c). In the MSE curve resulting from the random initial condition each $SE(\tau)$ value is slightly perturbed from the original, but the general behavior of the MSE curve remains unchanged. Thus, long range correlations and entropy are independent of the initial condition used for simulating $\mathcal{N}$.

The link between the complexity of a time series produced by a free-running physiological system and that system's robustness and health has potential applications in the diagnosis and treatment of physiological ailments. To make the leap to clinical applications, this observed correlation must be put on a firmer footing by understanding more precisely the causal link between the dynamics and structure of

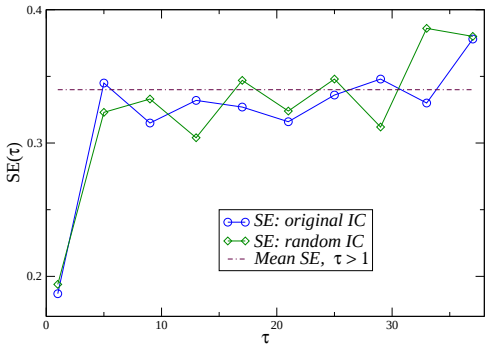


Figure 4.14: Comparison of MSE curves for the initial condition used in the simulation in Figures 4.9 and 4.11 and a randomly chosen initial condition for the excitatory cell e_1. For $\tau \geq 5$, $\overline{SE}(\tau) = 0.34$ for both curves. The original initial condition is -0.5; the randomly initial condition is 0.7957.

the system on one hand, and the time series structure on the other. Mathematical models will play a decisive role in this process, since they allow for direct testing of this connection.

It has proven quite challenging to construct such models. We report here on a successful attempt, where we show that a randomly connected small network of FitzHugh-Nagumo neurons can reproduce detrended fluctuations and multiscale entropy observed in physiological time series. In analyzing this model, we have found that when the length of the time series exceeds some critical length, the DFA and MSE measurements from the time series remain relatively constant, and, in addition, they do not depend on initial conditions. This confirms that the statistics computed from the time series reflect properties of the underlying system, rather than the particulars of the measurement process.

In addition, we have made two important observations. Firstly, given a neural network $\mathcal{N}$ and simulations of various lengths, the intrinsic complexity of the signal is maintained over a wide range of time series lengths. This is key for computations

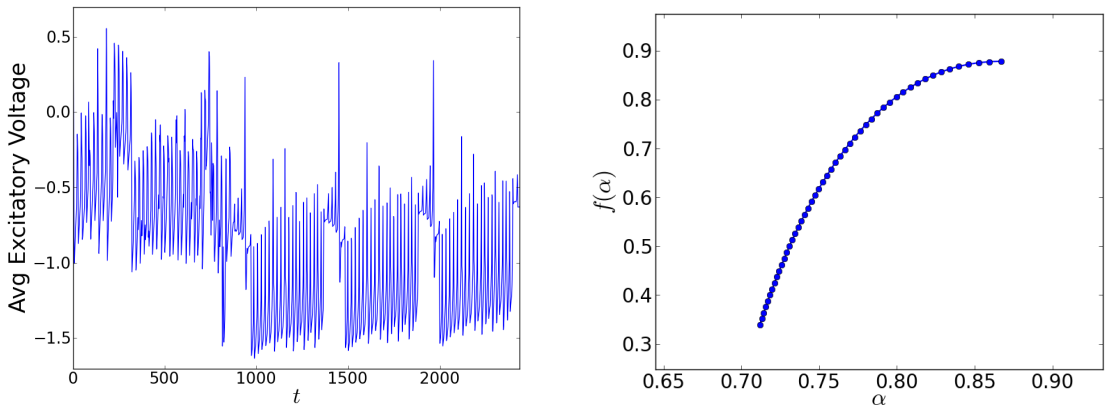


Figure 4.15: (a) The times series $g(t)$ generated by the average excitatory voltage of the network of neurons in Figure 4.8. (b) The multifractal spectrum for the density of $g(t)$ computed using the simplicial measure.

involving optimization of network topology and parameter sets. It allows us to run large batches of simulations, each for a relatively short amount of time, confident that physiological complexity seen in the resulting time series coincides with that of a longer series.

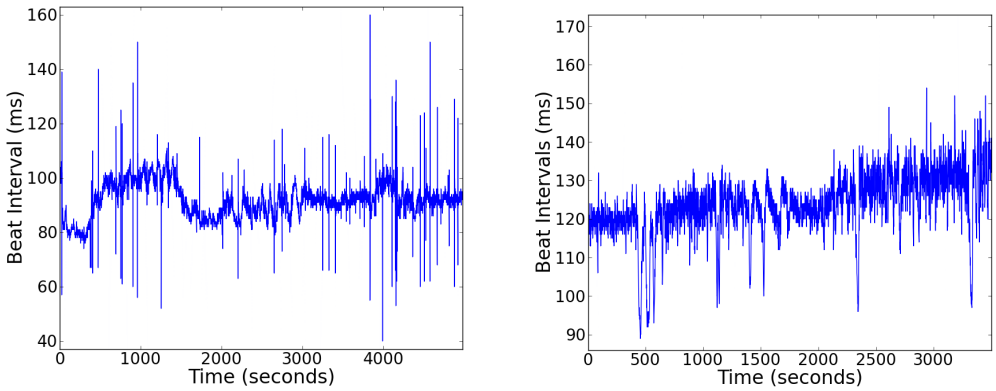
Secondly, there exists a finite range of time scales over which the network displays complex behavior. This shows that the network $\mathcal{N}$ has a distinct limit to its capacity for complexity and is incapable of producing complexity at every time scale. This result serves to clarify the boundaries of an *in-silico* neural network's complex behavior. Thus, we can determine the upper bound of complexity inherent to each network and optimize with respect to this measurement as well. Future work will focus on expanding this boundary by optimizing over network size and coupling strengths.

The ultimate test of our model, however, is its ability to match concrete experimental physiological data. We are currently collaborating with a group that has access to clinical data for both healthy and compromised individuals to see whether

our model can simulate the statistics of physiological measurements obtained from both of these groups.

4.4 Multifractal Analysis of Neural Networks

Physiological systems have been shown to exhibit a wide range of fractal behaviors [25, 37, 42, 49]. The results in 4.3 are aimed at developing a computational model that reproduces this behavior, and that can be probed and studied. The long-term goal in developing the simplicial measure and the computational techniques behind them is to analyze real-world dynamical systems and mathematical models of those systems. We have applied these techniques to the average excitatory voltage time series from Section 4.3. Call this time series $g(t)$.



(a) Healthy Individual (b) Compromised Individual

Figure 4.16: (a) and (b) show snapshots of inter-beat interval time series recorded for a patient with congestive heart failure and a healthy individual, respectively.

A closeup of $g(t)$ is shown in Figure 4.11(a). As mentioned above, we found that the series exhibits physiological complexity when analyzed with the DFA and MSE metrics (Figures 4.11(b) and 4.11(c)). Indeed, by using the simplicial measure to compute the $f(\alpha)$ spectrum for g, we see a relatively wide range of scaling exponents,

which has been shown to be indicative of a healthy, free-running system. This confirms the earlier results for the excitatory time series when analyzed using DFA and MSE. The scaling range of α is not as wide as the physiological system investigated in [49], but that is to be expected as the size of the system is far smaller.

4.5 Application to Physiological Data

We end this chapter with an analysis of human heart beat data. Recall that the health of a free-running physiological system has been correlated to the complexity of time series gathered from measuring the system. As noted in Section 4.3.1, many studies of physiological system have bolstered this hypothesis [26, 28, 41, 40, 66, 64]. We obtained data from PhysioBank [41], a database containing thousands of recording of numerous physiological processes. The following time series were studied:

i. RR intervals for a healthy, sleeping individual (Ironman contestants);

ii. RR intervals for an individual suffering from congestive heart failure.

The term "RR interval" refers to the the time interval, in milliseconds, between heart beats when analyzed with a traditional electrocardiogram (ECG). These are the gaps between the main peaks visible in an ECG. In what follows we have analyzed the density of the inter-beat interval signal. Examples of inter-beat interval graphs are plotted in Figures 4.16(b) and 4.16(a).

Recall from Section 2.4 on page 30 that the magnitude of the wavelet transform applied to a discrete data set is proportional to the density of points in a small region around the center of the wavelet. The modulus of the wavelet transform of a function f at scale a is proportional to the difference quotient of f over an interval of length a about the center of the wavelet. Thus, we can approximate this density of points in a

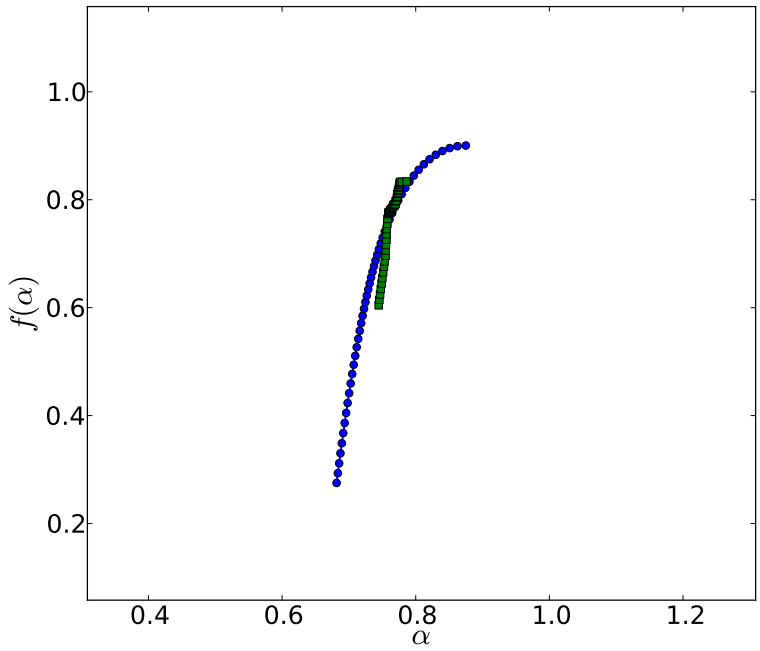


Figure 4.17: Multifractal spectra for a distribution derived from the time series in Figure 4.16 computed using the simplicial measure. The spectrum from a healthy individual (blue circles) produces a wider range of scaling exponents than the spectrum for the CHF patient (green squares), indicating greater complexity in the time series.

small interval by computing the absolute value of the difference quotient of f across this interval. In this way, we approximate a distribution.

For the individual reporting congestive heart failure, we compute the $f(\alpha)$ spectrum (green squares) in Figure 4.17, which clearly shows a very narrow scaling range. The scaling exponents α range from approximately 0.74 to 0.78. This result agrees with previous results for wavelets applied to inter-beat time series reported by Goldberger, *et al.* [42] and Ivanov [49]. Alternatively, the spectrum for the healthy individual's inter-beat recording in Figure 4.17 (blue circles), contains a far wider range of scaling exponents. In this case, we find that α ranges from approximately 0.65 to 0.85. The width of the ranges for α found in both cases agree with the ranges seen for Hölder exponents when using wavelet analysis in results of Ivanov, *et al.*. [49].

4.6 Summary of Results

We study dynamical systems of varying complexity in the sections above, utilizing the simplicial measure as the basis for approximating the SBR measure on the attractor. In the case of measure concentrated on the Cantor set, the approximation of the multifractal spectrum is nearly exact (for $q > 0$). For both the Cantor set and Hénon attractor, the simplicial measure efficiently computes the fractal dimension of the support of the SBR measure with a high degree of accuracy. In the case of the Hénon attractor, we note two important points. First, there is a shift in the scaling exponent. This is likely due to numerical error. Nevertheless, the computation of the dimension of the attractor ($f(q)$) is still highly accurate–and our method uses approximately five orders of magnitude fewer points to compute the $f(\alpha)$ spectrum than previous experiments.

The long-term thrust of this work, though, is towards applications to mathematical models as well as the analysis of real-world data. In Sections 4.4 and 4.5 we have demonstrated that the simplicial measure can be used successfully to study these more complex systems. Working with a neural network that has been shown to produce time series that are physiologically complex, we are able to confirm these physiological characteristics by studying the range of scaling exponents in the $f(\alpha)$ curve. Bolstering these results are the analyses of heart beat time series. Again, previous results have shown a marked decrease in the range of the scaling exponents α in the computation of the $f(\alpha)$ curve for time series obtained from individuals with compromised health. The comparison of the curves in Figure 4.17, computed using our methods, shows definitively that we are able to differentiate time series from healthy and compromised individuals based on multifractal analysis using simplicial measures.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

Investigation of complex natural phenomena uses increasingly sophisticated mathematical tools. Numerous methods have been developed to understand and quantify trends in data collected from complex, real-world systems, as well as dynamical systems modeling such systems. Often as important is an understanding of system's qualitative behavior. In this manuscript we have developed and applied a new technique to distinguish qualitative invariants of dynamical systems. We have concerned ourselves with questions focusing on what can be learned from the observation of, and the acquisition of data from, a dynamical system. In particular, we have studied the Sinai-Bowen-Ruelle (SBR) measure on an attractor of a dynamical system.

The study of the SBR measure μ on an attractor A of a dynamical system f makes use the ergodic averages,

$$\mu = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \delta_{f^i(x)},$$

discussed in Section 2.1. We have approached this problem from a new direction, constructing approximations of the measure μ in a local manner. Specifically, given a finite data set $Y = \{x, f(x), f^2(x), \ldots, f^N(x)\}$ and a scale ε, we have approximated the density of the mass distribution on A by constructing a neighborhood graph $\mathcal{N}_\varepsilon(Y)$ on Y. We have shown that the extension of the neighborhood graph to a Vietoris-Rips complex, $\mathcal{R}_\varepsilon(Y)$, contains within its maximal simplices an approximation of the local density at scale ε. We termed this the simplicial measure ν.

We then demonstrated, for μ concentrated on a multiplicative set in one dimension, that μ is approximated to arbitrary precision by ν. Hence, we can utilize the simplicial measure to approximate the multifractal spectrum of μ. In addition, there is a rich tradition of using wavelets to analyze fractal signals, and we have shown that the

simplicial measure agrees with the Haar wavelet transform of μ on such multiplicative sets.

The original purpose in developing the simplicial measure was to study complex systems. In order to show the efficacy of the simplicial measure we used it to approximate the measure on a number of data sets. We first considered precisely constructed mathematical sets. In one dimension, these sets included approximations of both mono- and multifractal measures supported on the Cantor set. The multifractal spectrum computed using the simplicial measure was demonstrated to be impressively accurate for non-negative values of q in the partition function.

In two dimensions we have studied the Hénon map, which is a common test piece for computational methods in dynamical systems. Our methods do not exhibit the same range as seen in previous work [6] of the scaling exponent α in the $f(\alpha)$ multifractal spectrum. Nevertheless, we attain very good results for the range of $f(\alpha)$, and an accurate approximation of the dimension of the support of the measure, using at least four orders of magnitude fewer points.

Time series obtained from healthy versus compromised physiological systems have previously been shown to exhibit qualitative differences. In [17], we produced by a mathematical model of a neural network that showed characteristics of a healthy physiological system, according to metrics utilized by specialists in the field [30, 42, 50, 64]. In Section 4.4 we studied the density distribution computed from the time series of the average voltage potential of the excitatory cells. We found that the range of scales for α, as well as the dimension for $f(\alpha)$, in the multifractal spectrum is indicative of healthy physiological systems, using [50] as a comparison.

To apply our methods to real-world data, we chose to use RR intervals for heart beats as explained in Section 4.5. As noted in [49, 50], the range of scales seen in plots of the multifractal spectra vary greatly depending the health of the individual.

Again studying the distribution determined from the time series, we applied the simplicial measure technique to the computation of the multifractal spectra. Our results showed a wide range of scaling exponents in the multifractal spectrum computed from the time series of a healthy individual, while the compromised individuals showed a markedly smaller range of scaling exponents. The size of the ranges for α agree with those of Ivanov, *et al.* [50]. It should be noted, though, that the values for α computed using wavelets in Ivanov, *et al.*, are shifted from those computed in Chapter 4. This area of research is still young, and we are currently investigating these questions.

5.1 Future Work

Finally, we would like to indicate paths that we plan to follow for further research. The results that we obtained for the Hénon attractor, while impressive, left room for improvement. We would like to investigate ways to bring the range of α closer to the range seen in the earlier work of Arneodo, *et al.*. Also along this path, we would like to guarantee that a comparably small number of points are used in extending the range of α. This path is computational in nature, and will involve a refinement of the current Vietoris-Rips construction algorithm.

The core of this dissertation, though, sits at the intersection of complex, physical systems and computational methods of dynamical systems. Thus, we plan to study further the physiological results described in Section 4.5, as well as extend the techniques to other complex systems, both in the real-world and modeled. After all, questions about complex systems from a wide range of fields deserve our full attention as scientists and mathematicians.

REFERENCES CITED

[1] Patrice Abry, Paulo Goncalves, and Jacques Levy Vehel (eds.), *Scaling, Fractals, and Wavelets*, Wiley & Sons, Hoboken, New Jersey, 2009.

[2] A Arneodo, E Bacry, and J Muzy, *The thermodynamics of fractals revisited with wavelets*, Physica A: Statistical and Theoretical Physics **213** (1995), no. 1-2, 232–275.

[3] A. Arneodo, E. Bacry, and J. F. Muzy, *Random cascades on wavelet dyadic trees*, Journal of Mathematical Physics **39** (1998), no. 8, 4142.

[4] A Arneodo, G. Grasseau, and M. Holschneider, *Wavelet transform of multifractals*, Physical review letters (1988).

[5] A. Arneodo, G. Grasseau, and M. Holschneider, *Wavelet transform of multifractals*, Physical review letters **61** (1988), no. 20, 2281–2284.

[6] A Arneodo, G Grasseau, and J. Kostelich, *fractal dimension and f(\alpha) spectrum of the henon attractor*, Physics Letters A **124** (1987), no. 8.

[7] J. Arrault, *Wavelet Based Multifractal Analysis of Rough Surfaces: Application to Cloud Models and Satellite Data*, Physical Review Letters **59** (1997), no. 1, 1424–78.

[8] Sunil Arya, David M. Mount, Nathan S. Netanyahu, Ruth Silverman, and Angela Y. Wu, *An optimal algorithm for approximate nearest neighbor searching fixed dimensions*, Journal of the ACM **45** (1998), no. 6, 891–923.

[9] Sunil Arya and D.M. Mount, *Approximate nearest neighbor queries in fixed dimensions*, Proceedings of the fourth annual ACM-SIAM Symposium on Discrete algorithms, Society for Industrial and Applied Mathematics, 1993, pp. 271–280.

[10] E. Bacry, J. F. Muzy, and A. Arneodo, *Singularity spectrum of fractal signals from wavelet analysis : exact results*, Journal of statistical physics **70** (1993), no. 3-4, 635–674 (eng).

[11] R Badii, *Renyi dimensions from local expansion rates*, Physical Review A **35** (1987), no. 3, 1288–1293.

[12] M.S. Baptista, Celso Grebogi, and E. Barreto, *Topology of Windows in the High-Dimensional Parameter Space of Chaotic Maps*, International Journal of Bifurcation and Chaos **13** (2003), no. 9, 2681–2688.

[13] L.M. Barreira, *Cantor sets with complicated geometry and modeled by general symbolic dynamics*, Random and Computational Dynamics **3** (1995), no. 4, 213–240.

[14] Ernest Barreto, Brian Hunt, Celso Grebogi, and James Yorke, *From High Dimensional Chaos to Stable Periodic Orbits: The Structure of Parameter Space*, Physical Review Letters **78** (1997), no. 24, 4561–4564.

[15] Robert G. Bartle, *The Elements of Integration and Lebesgue Measure*, John Wiley & Sons, New York, NY, 1995.

[16] Jesse Berwald, *simplicial-measures*, 2011.

[17] Jesse Berwald, Tomas Gedeon, and Konstantin Mischaikow, *Modeling Complexity of Physiological Time Series In-Silico*, BioSignals Proceedings, 2010.

[18] Patrick Billingsley, *Ergodic Theory and Information*, 1st ed., John Wiley & Sons, New York, 1965.

[19] Rufus Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, 2nd ed., Springer-Verlag, Berlin, 2008.

[20] C Bron and J Kerbosch, *Algorithm 457: Finding All Cliques of an Undirected Graph*, Communications of the ACM **16** (1973), no. 9, 575–577.

[21] E. Carlsson, Gunnar Carlsson, V. De Silva, and SJ Fortune, *An algebraic topological method for feature identification*, International Journal of Computational Geometry and Applications **16** (2006), no. 4, 291–314.

[22] F Cazals and C. Karande, *A note on the problem of reporting maximal cliques*, Theoretical Computer Science **407** (2008), no. 1-3, 564–568.

[23] A. Chhabra and R.V. Jensen, *Direct determination of the $f(\backslash alpha)$ singularity spectrum*, Physical Review Letters **62** (1989), no. 12, 1327–1330.

[24] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, *Introduction to Algorithms*, third ed., MIT Press, Cambridge, MA, 2009.

[25] M. Costa, Ionita Ghiran, CK Peng, A Nicholson-Weller, and A.L. Goldberger, *Complex dynamics of human red blood cell flickering: Alterations with in vivo aging*, Physical Review E (2008), 1–4.

[26] Madalena Costa, Ary Goldberger, and C.-K. Peng, *Multiscale Entropy Analysis of Complex Physiologic Time Series*, Physical Review Letters **89** (2002), no. 6, 6–9.

[27] ———, *Multiscale Entropy Analysis of Complex Physiologic Time Series*, Physical review letters **89** (2002), 068102.

[28] ———, *Multiscale entropy analysis of biological signals*, Physical Review E **71** (2005), no. 2, 1–18.

[29] Madalena Costa, C-K Peng, and Ary L Goldberger, *Multiscale Analysis of Heart Rate Dynamics: Entropy and Time Irreversibility Measures.*, Cardiovascular engineering (Dordrecht, Netherlands) **8** (2008), no. 2, 88–93.

[30] Madalena D Costa, Chung-Kang Peng, and Ary L Goldberger, *Multiscale analysis of heart rate dynamics: entropy and time irreversibility measures.*, Cardiovascular engineering (Dordrecht, Netherlands) **8** (2008), no. 2, 88–93.

[31] Ingrid Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

[32] V. de Silva and R Ghrist, *Coverage in sensor networks via persistent homology*, Algebraic & Geometric Topology **7** (2007), 339–358.

[33] J.-P. Eckmann and D. Ruelle, *Ergodic theory of chaos and strange attractors*, Reviews of modern physics **57** (1985), no. 3.

[34] Kenneth Falconer, *Geometry of Fractal Set*, Cambridge University Press, Cambridge, 1985.

[35] ———, *Techniques in Fractal Geometry*, 1st ed., Wiley & Sons, West Sussix, England, 1997.

[36] ———, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley & Sons, West Sussix, England, 2003.

[37] J. Gao, J. Hu, W. Tung, and Y. Cao, *Distinguishing chaos from noise by scale-dependent Lyapunov exponent*, Physical Review E **74** (2006), no. 6, 66204.

[38] J. Ghez and Others, *Integrated wavelets on fractal sets. I. The correlation dimension*, Nonlinearity **5** (1992), 777.

[39] Robert Ghrist, *Barcodes: The persistent topology of data*, Bulletin of the American Mathematical Society **45** (2008), no. 1, 61.

[40] Ary L Goldberger, *Giles F. Filley lecture. Complex systems.*, Proceedings of the American Thoracic Society **3** (2006), no. 6, 467–71.

[41] Ary L Goldberger, L.A.N. Amaral, L Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, RG Mark, JE Mietus, GB Moody, C-K Peng, and H.E. Stanley, *PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals*, Circulation **101** (2000), no. 23, e215–e220.

[42] Ary L Goldberger, Luis A. N. Amaral, Jeffrey M Hausdorff, Plamen Ch Ivanov, C-K Peng, and H Eugene Stanley, *Fractal dynamics in physiology: alterations with disease and aging*, Proceedings of the National Academy of Sciences of the United States of America **99 Suppl 1** (2002), 2466–72.

[43] Peter Grassberger, *Generalized dimensions of strange attractors*, Physics Letters A **97** (1983), no. 6, 227–230.

[44] Alfred Haar, *Theorie der orthogonalen Funktionensysteme*, Mathematische Annalen **69** (1910), no. 3, 331–371.

[45] Allen Hatcher, *Algebraic Topology*, Cambridge University Press, Cambridge, MA, 2002.

[46] M. Hénon, *A two-dimensional mapping with a strange attractor*, Communications in Mathematical Physics **50** (1976), no. 1, 69–77.

[47] John G. Hocking and Gail S. Young, *Topology*, Dover, New York, 1961.

[48] Jing Hu, Jianbo Gao, and Xingsong Wang, *Multifractal analysis of sunspot time series: the effects of the 11-year cycle and Fourier truncation*, Journal of Statistical Mechanics: Theory and Experiment **2009** (2009), P02066.

[49] P C Ivanov, L A Amaral, A L Goldberger, S Havlin, M G Rosenblum, Z R Struzik, and H E Stanley, *Multifractality in human heartbeat dynamics.*, Nature **399** (1999), no. 6735, 461–5.

[50] Plamen Ch. Ivanov, Luis a. Nunes Amaral, Ary L. Goldberger, Shlomo Havlin, Michael G. Rosenblum, H. Eugene Stanley, and Zbigniew R. Struzik, *From 1/f noise to multifractal cascades in heartbeat dynamics.*, Chaos (Woodbury, N.Y.) **11** (2001), no. 3, 641–652.

[51] Tomasz Kaczynski, Konstantin Mischaikow, and Marian Mrozek, *Computational Homology*, Springer, 2004.

[52] Anatole Katok and Boris Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, 1995.

[53] Dmitry Kozlov, *Combinatorial Algebraic Topology*, Springer, Berlin, 2008.

[54] Kapilanjan Krishan, Huseyin Kurtuldu, Michael F. Schatz, Marcio Gameiro, Konstantin Mischaikow, and Santiago Madruga, *Homology and symmetry breaking in Rayleigh-Benard convection: Experiments and simulations*, Physics of Fluids **19** (2007), no. 11, 117105.

[55] Henrik Madsen, *Time Series Analysis*, 1st ed., Chapman & Hall, Boca Raton, FL, 2008.

[56] S. Mallat and W.L. Hwang, *Singularity detection and processing with wavelets*, IEEE Transactions on Information Theory **38** (1992), no. 2, 617–643.

[57] B. Mandelbrot, *How long is the coast of Britain? Statistical self-similarity and fractional dimension*, Science **156** (1967), no. 3775, 636.

[58] Benoit B. Mandelbrot, *Fractal Geometry of Nature*, W.H. Freeman, 1983.

[59] C. Meneveau and KR Sreenivasan, *Simple multifractal cascade model for fully developed turbulence*, Physical review letters **59** (1987), no. 13, 1424–1427.

[60] James R. Munkres, *Elements of Algebraic Topology*, Addison-Wesley, Redwood City, CA, 1984.

[61] JF Muzy, E. Bacry, and A. Arneodo, *Wavelets and multifractal formalism for singular signals: application to turbulence data*, Physical Review Letters **67** (1991), no. 25, 3515–3518.

[62] ———, *Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method*, Physical review E **47** (1993), no. 2, 875–884.

[63] C.-K. Peng, SV Buldyrev, S. Havlin, M. Simons, HE Stanley, and AL Goldberger, *Mosaic organization of DNA nucleotides*, Physical Review E **49** (1994), no. 2, 1685–1689.

[64] C.-K. Peng, Shlomo Havlin, H.E. Stanley, and A.L. Goldberger, *Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series*, Chaos An Interdisciplinary Journal of Nonlinear Science **5** (1995), no. 1, 82.

[65] C-K Peng, Albert C-C Yang, and Ary L Goldberger, *Statistical physics approach to categorize biologic signals: from heart rate dynamics to DNA sequences.*, Chaos (Woodbury, N.Y.) **17** (2007), no. 1, 015115.

[66] C.K. Peng, SV Buldyrev, S. Havlin, M. Simons, HE Stanley, and AL Goldberger, *Mosaic organization of DNA nucleotides*, Physical Review E **49** (1994), no. 2, 1685–1689.

[67] Yakov B. Pesin, *Dimension Theory in Dynamical Systems: Contemporary views and applications*, University of Chicago Press, Chicago, 1997.

[68] J S Richman and J R Moorman, *Physiological time-series analysis using approximate entropy and sample entropy.*, American journal of physiology. Heart and circulatory physiology **278** (2000), no. 6, H2039–49.

[69] SG Roux, A. Arneodo, and N. Decoster, *A wavelet-based method for multifractal image analysis. III. Applications to high-resolution satellite images of cloud structure*, The European Physical Journal B **15** (2000), no. 4, 765–786.

[70] David Ruelle, *A Measure Associated with Axiom-A Attractors*, American Journal of Mathematics **98** (1976), no. 3, 619–654.

[71] Marten Scheffer, Jordi Bascompte, W.A. Brock, Victor Brovkin, S.R. Carpenter, Vasilis Dakos, Hermann Held, E.H. Van Nes, Max Rietkerk, and George Sugihara, *Early-warning signals for critical transitions*, Nature **461** (2009), no. 7260, 53–59.

[72] Ruedi Stoop, Philipp Benner, and Yoko Uwate, *Real-World Existence and Origins of the Spiral Organization of Shrimp-Shaped Domains*, Physical Review Letters **105** (2010), no. 7, 2–5.

[73] F Takens, *Detecting strange attractors in turbulence*, Dynamical systems and turbulence, Warwick 1980 (1981), 366–381.

[74] D Terman, S Ahn, X Wang, and W Just, *Reducing neuronal networks to discrete dynamics*, Physica D: Nonlinear Phenomena **237** (2008), no. 3, 324–338.

[75] K.W. Thoning, P.P. Tans, W.D. Komhyr, and Others, *Atmospheric carbon dioxide at Mauna Loa Observatory, 2, Analysis of the NOAA/GMCC data, 1974-1985*, J. Geophys. Res **94** (1989), 8549–8565.

[76] Peter Walters, *An Introduction to Ergodic Theory*, Springer, New York, NY, 1982.

[77] A Zomorodian, *Fast construction of the Vietoris-Rips complex*, Computers & Graphics (2010).

APPENDICES

APPENDIX A

PARAMETERS USED TO GENERATE RESULTS FOR NEURAL NETWORK SIMULATIONS

The network $\mathcal{N}$ used in Section 4.3.8 is generated by the 10-cell graph in Figure 4.8 with the parameter set P listed in the following tables.

Table A.1: Parameter set for neural network $\mathcal{N}$, excluding ε's

α	α_I	α_x	β	β_I	β_x	g_{EI}	g_{IE}	g_{II}
4.0	4.0	1.0	0.1	0.1	4.0	0.4	0.4	0.4
v_I	v_E	θ	θ_I	θ_x	b	c	K_I	K_E
3.0	0.1	0.1	0.1	0.1	0.8	0.7	0.28	0.35

Table A.2: ε set for $\mathcal{N}$

Excitatory	ε_1	0.08456607
	ε_2	0.00043158
	ε_3	0.00068327
	ε_4	0.06293498
	ε_5	0.00537958
Inhibitory	ε_1	0.00017724
	ε_2	0.03678080
	ε_3	0.05379177
	ε_4	0.00140943
	ε_5	0.00037465