FRACTAL ANALYSIS

IN

A ONE-DIMENSIONAL UNIVERSE

by

YUI SHIOZAWA

Bachelor of Science, 2010
University of Toronto
Ontario, Canada

Submitted to the Graduate Faculty of the
College of Science and Engineering
Texas Christian University
in partial fulfillment of the requirements
for the degree of

Master of Science

December 2014

FRACTAL ANALYSIS IN A ONE-DIMENSIONAL UNIVERSE

by

Yui Shiozawa

Thesis approved:
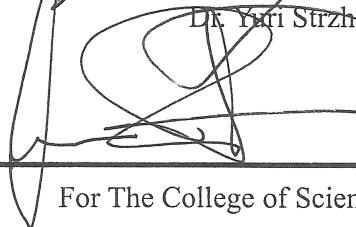
Major Professor: Dr. Bruce Miller

Dr. Hana Dobrovolny

Dr. Peter Frinchaboy

Dr. Igor Prokhorenkov

Dr. Yuri Strzhemechny

For The College of Science and Engineering

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my research advisor, Dr. Bruce N. Miller for his unfaltering support, guidance and expertise. Next, I would like to thank our collaborator Dr. Jean-Louis Rouet for making this research possible. I also benefitted from Dr. Igor Prokhorenkov's insightful comments and advice. Additionally, I would like to thank my committee members. Finally, I would like to acknowledge all the faculty and graduate students in the Physics and Astronomy Department for their generous support both in and outside the classroom.

Contents

List of Figures

Abbreviations

CDM	**C**old **D**ark **M**atter
CMB	**C**osmic **M**icrowave **B**ackground
DF	**D**imension **F**unction
KS	**K**olmogorov **S**mirnov
MBP	**M**ultiplicative **B**inomial **P**rocess
RF	**R**ouet **F**elix
SDSS	**S**loan **D**igital **S**ky **S**urvey
WMAP	**W**ilkinson **M**icrowave **A**nisotropy **P**robe

Chapter 1

Introduction

Fractals are mathematical sets characterized by self-similarity. While the history of the study of fractals goes back as far as the 17th century [1], the concept was popularized by Mandelbrot in the 1970s [2] and is now applied to many fields including cosmology [3], geography [4], chemistry [5], and economics [6]. The fact that fractals can be found virtually everywhere suggests that there is an underlying mathematical principle. From the onset of fractal studies, some have suggested that the universe might be a fractal at a large scale [7].

The cosmological principle holds that the universe is homogeneous and isotropic. Indeed, the observed cosmic microwave background, or simply CMB, which is believed to be a relic from the primordial universe, suggests that the universe was highly homogenous and isotropic at its early stage. The cosmological principle, though appealing from a naturalistic perspective as well, is nevertheless a hypothesis. In fact, in the past few decades, as we probed further into the dark, we encountered ever larger structures. The Laniakea supercluster, one of the largest structures in the universe, was discovered in 2014 [8], further redefining the structure horizon.

The standard theory contends that, in the critical universe, the slightly over-dense regions act as seeds for gravitationally bound structures, evolving into galaxies and galactic clusters with large voids between them. This bottom-up structure formation also indicates that the universe is homogenous and isotropic beyond a certain scale range, supporting the cosmological principle without undermining what we observe today.

While extensive research efforts have been devoted to simulating possible scenarios, most of which closely outline the standard model, the specifics of how the large structure has been formed from little-to-no structure remains unresolved. Within the standard scenario, the scale at which the universe is safely considered homogenous is disputed among researchers. In order to replicate how the universe evolved, three-dimensional simulations are frequently employed. However, these three dimensional simulations are not suitable for fractal analysis for a number of reasons [9]: Fractal analysis typically requires a large number of sample points for accurate results. Moreover, in order to investigate the fractal features of a given set, a sample set needs to properly reflect the fine structure of that set. Three-dimensional simulations to date have been performed with a limited number of particles per dimension and often make numerical approximations in compensation of computational time and simplicity. These approximations may destroy the fine structure that the simulated set may possess and introduce unknown effects into the evolution of the universe.

With a focus on fractal analysis, we have developed a one-dimensional analogue of the universe that allows us to write an event-driven algorithm for a large number of matter particles. In this variant, we follow the evolution of two kinds of matter, namely dark matter and luminous matter. It is largely accepted that the ordinary matter that is visible to us, therefore is luminous, only accounts for roughly 20% of the matter content in the universe. The remainder is referred to as dark matter, which interacts with other matter only via gravity. Accordingly, we simulated luminous matter in a way that it loses energy in interaction while dark matter distribution merely evolves under the influence of gravity.

The preliminary results of our simulations clearly show the development of structure formation that exhibits characteristic fractal patterns over a certain scale range. In particular, the structure appears to be hierarchical, similar to the observational universe which also appears to be arranged in a hierarchical pattern, from a galaxy group to a supercluster. This hierarchical structure is the hallmark of fractal geometry and is characterized by self-similarity. Therefore, to study the structure created in the simulation, we focused our attention in fractal analysis.

Central to fractal geometry is the concept of self-similarity. Due to its scale-free properties, finding appropriate power laws plays a crucial role in fractal analysis. By extending the traditional notion of dimension, fractal dimensions may be extracted from the associated exponent of these power laws. For instance, the box-counting dimension D_0, one of the fractal dimensions, roughly represents the "size" of a given set. If the matter constituents of the universe are homogeneously distributed, the dimension should be close to the dimension of the embedded space. According to the earliest computational results, the correlation dimension D_2, also one of the fractal dimensions, was estimated to be about $D_2 \simeq 1.2$ [10].

For dynamical systems, multifractals often arise, requiring a spectrum of fractal dimensions, called the generalized dimensions, instead of a single dimension. A multifractal analysis allows us to investigate the subsets of the original set with various densities. Martinez *et al.* performed a multifractal analysis, obtaining the spectrum of dimension D_q [11]. They found that although the scaling range is not trivial, the distribution of galaxies exhibits highly multifractal characteristics. They estimated box-counting dimension $D_0 \simeq 2.1$ and correlation dimension $D_2 \simeq= 1.3$. The fact that these computed fractal dimensions were significantly smaller that the embedded dimension indicates that the universe is a fractal in a certain range.

Although size-oriented methods are most popular for computing the generalized dimensions, they are known to be unreliable in a certain range of the spectrum. Accordingly, we pursued two alternative methods, both of which employ mass-oriented partitions. One of these methods, namely the k-neighbor method, proves suitable for analyzing structure formation. From this fractal analysis technique, we justify the aforementioned claim that our simulations show the development of structure formation up to a certain scale.

This thesis is organized as follows: Chapter 2 introduces the concepts and definitions regarding fractals; Chapter 3 discusses the specific methods for computing fractal dimensions. Most of the details can be found in our published work [12]; Chapter 4 includes the basic cosmology required to construct the gravitational simulation; Chapter 5 explains the one-dimensional model, presents the preliminary results of the simulation and discusses future directions.

Chapter 2

Fractal Analysis

2.1   Fractals

The discovery of a rather strange type of set, called fractal, led to the extension of the notion of dimension. Fractals sets are characterized by self-similarity, and power laws can be associated with them. From a geometrical perspective, a given set is self-similar when it is similar to a part of itself, though the rigorous definitions of fractals or self-similarity are often avoided in order to encompass a wider variety of fractals. To achieve the self-similarity condition, a self-similar set needs to possess an infinite nesting structure. Due to this, fractals may be preserved after appropriate magnification and translation. Therefore, power laws arise naturally in the study of fractals as the power law is the only differentiable function that does not change its form under a scale transformation. To be precise, if some differentiable function f satisfies $f(bx) = g(b)f(x)$ for all $b > 0$ and for some function g, the function f must be a power law. Since $x \rightarrow bx$ is a scale transformation, the function f is said to be preserved up to a constant under scale transformations. Accordingly, various power laws can be derived from fractal sets and it is the exponents of these power laws that are associated with the dimensions of the fractal set.

Traditionally, the dimension of a given set indicates the number of independent variables required to specify the element within the set and so can take only integer values.

However, as in the case of space-filling curves, this notion of dimension may not be satisfactory. Instead, fractal dimension is better suited to represent the "size" of a given set. If we want to associate "size" with fractals such as the famous Koch snowflake [13] shown in Fig. 2.1, we need to extend our notion of dimension as well as of measure. The Koch snowflake is nowhere differentiable and consists of a perimeter with infinite length enclosing a finite area. Intuitively, the dimension of the set should be bigger than the dimension of a finite interval and smaller than that of a finite area. Indeed, we can define the fractal dimension in such a way that the Koch snowflake has the dimension of $\log 4 / \log 3 = 1.261....$ In this section, we introduce all the necessary concepts to define fractal dimension.

FIGURE 2.1: **The Koch Snowflake is illustrated at the third iteration. The Koch Snowflake can be created as follows: Starting with an equilateral triangle, place another equilateral triangle on each of the sides of this original triangle. Repeat this process indefinitely on the straight lines newly formed at each iteration. A simple (finite) self-similarity can be seen in this illustration.**

2.1.1 Measure

The classical notion of dimension is, roughly, the number of independent variables required to specify a point in a given set. Here, we extend the classical notion of dimension by using the concept of measure (see Appendix A for a more detailed discussion of measure). The most standard and robust definition of dimension is based on the Hausdorff measure. Unlike the familiar Lebesgue measure, it is defined for all sets. Suppose F is a subset of $\mathbb{R}^n$ and s is a non-negative number. For any δ, define [14]

$$\mathscr{H}_\delta^s(F) := \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ is a } \delta\text{-cover of } F \right\} \tag{2.1}$$

(see Appendix A.2 for the definition of a cover). Then the s-dimensional Hausdorff measure of F is given by taking the limit as $\delta \to 0$:

$$\mathscr{H}^s(F) = \lim_{\delta \to 0} \mathscr{H}^s_\delta(F). \tag{2.2}$$

It can be shown that when applied to "simple" n-dimensional sets, the n-dimensional Hausdorff measure is consistent with our classical notion of n-dimensional volume of the set, namely length, area etc. In fact, it can be shown that if F is a Borel subset of $\mathbb{R}^n$, then the Hausdorff measure coincides with the Lesbegue measure up to a constant multiple.

2.1.2 Dimension

Since $\mathscr{H}^s_\delta(F)$ in Eq. (2.1), and therefore the Hausdorff measure $\mathscr{H}^s(F)$ is a non-increasing function of s for $\delta < 1$, using the same notation, it can be shown that for some $t > s$, the following inequality holds [14]:

$$\sum_i |U_i|^t \le \delta^{(t-s)} \sum_i |U_i|^s. \tag{2.3}$$

This means that if $\mathscr{H}^s(F) < \infty$ for some s, $\mathscr{H}^t(F) = 0$ for all $t > s$. This property guarantees the existence of the critical value of s at which $\mathscr{H}^s(F)$ jumps from ∞ to 0. The Hausdorff dimension $\dim_\mathrm{H} F$ of a set F is defined to be this critical value. More formally, it can be written as [14]:

$$\dim_\mathrm{H} F = \inf \left\{ s : \mathscr{H}^s(F) = 0 \right\} = \sup \left\{ s : \mathscr{H}^s(F) = \infty \right\}. \tag{2.4}$$

Note that the concept of dimension has now been extended to non-integer values as $\dim_\mathrm{H} F$ can assume any real number. While the Hausdorff dimension might not be the most intuitive way to define dimension, it is consistent with the classical definition of integer dimension and also satisfies a set of properties that we expect to hold true, based on our geometrical notion, for any reasonable definition of dimension. For example, the Hausdorff dimension is monotonic as it should be: If $E \subset F$, then $\dim_\mathrm{H} E \le \dim_\mathrm{H} F$. It is this robustness of the

Hausdorff dimension that makes it widely regarded as the most foundational definition of dimension when it comes to fractal analysis.

2.1.3 Box-counting Dimension

While the Hausdorff dimension is mathematically rigorous, it is difficult to estimate directly. In many cases, the box-counting dimension, although less rigorous, is often used instead. In Eq. (2.1), consider a cover with $|U_i| = \delta$ for all i as illustrated in Fig. 2.2. If $n(\delta)$ denotes the number of elements of the cover required to cover a given set, we have [14]:

$$\mathscr{H}_\delta^s(F) = \inf_{|U_i| \leq \delta} \left\{ \sum_{i=1}^{\infty} |U_i|^s \right\} \leq n(\delta)\delta^s. \tag{2.5}$$

If we assume that the following scaling law holds as $\delta \to 0$:

$$n(\delta) \simeq \delta^{-D_0} \tag{2.6}$$

for some $D_0 \in \mathbb{R}_+$, we see that the R.H.S of Eq. (2.5) converges for $s \geq D_0$ as $\delta \to 0$. Therefore, we obtain the following relation:

$$D_H \leq D_0. \tag{2.7}$$

The exponent in Eq. (2.6), D_0, is called the box-counting dimension and exists for a variety of sets. The box-counting dimension has practical values as it can be computed as the slope of a log-log plot of Eq. (2.6). Note that the power law Eq. (2.6) may not exist and the box-counting dimension may be different from the Hausdorff dimension. Therefore, it is important to understand the limitation of the definition and its relation to the Hausdorff dimension. For example, a set of rational numbers has the Hausdorff dimension of zero and the box-counting dimension of one. Furthermore, the exponent D_0 may not converge to a single value. Yet, the box-counting dimension estimates the Hausdorff dimension very well for many physical sets found in nature.
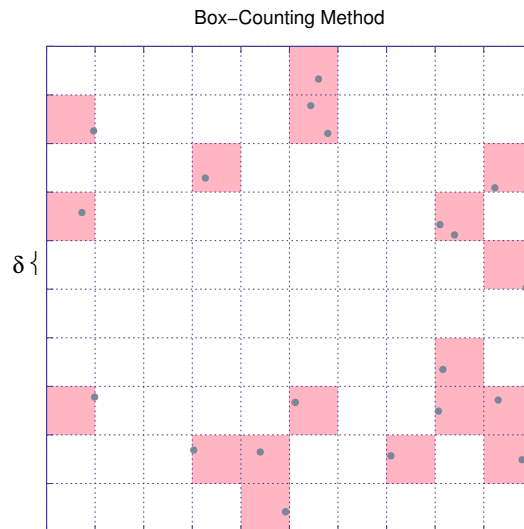
FIGURE 2.2: **A heuristic illustration of the box-counting method. The shaded "boxes" indicate the elements of a partition that contain the members of a given set represented by dots in the figure. Here, the box-size was set to δ.**

2.1.4 The Generalized Dimension

Behind the box-counting dimension was the key power law Eq. (2.6). However, this is not the only power law we can find with a given set. For instance, the average "information" required to identify a box containing the member of a given set also assumes the form of a power law as the box size goes to zero. This information is measured using the Shanon entropy. The associated scaling exponent is called the information dimension and is typically denoted by D_1. Moreover, the correlation dimension D_2 is often used due to its computational robustness. The correlation integral measures the average probability of a pair of sample points being close together and is the integral form of the correlation function. It turns out that the correlation integral takes the form of a power law as the number of sample points from a given set goes to infinity. Thus the correlation dimension D_2 can be extracted as the exponent of this power law. These exponents D_1 or D_2 are generally different from D_0 and each of them conveys an important piece of information about the set.

The generalized dimension, also known as Rényi dimension, generalizes these dimensions and extends the traditional notion of dimension even further by generating a spectrum of dimension instead of associating a single value to a given set. Suppose $C = \{U_i\}$ is a cover of a set $A \subset \mathbb{R}^n$. Let n_i denote the number of points in U_i among n randomly

chosen points from A. Then p_i is associated with U_i for each i by $p_i = \lim_{n \to \infty} \frac{n_i}{n}$. For any real number $q \neq 1$, the generalized dimension D_q for a set A is given by [14]:

$$D_q = -\frac{1}{1-q} \lim_{\epsilon \to 0} \frac{\ln \sum_{i=1}^{N(\epsilon)} p_i^q}{\ln \epsilon} \qquad (2.8)$$

where $N(\epsilon)$ is the number of sets with diameter $|U_i| = \epsilon$ required to cover the set A. For $q = 1$, the limiting case where $q \to 1$ is used. This notation is consistent with the earlier convention for the box-counting dimension and other dimensions. For instance, the box-counting dimension D_0, D_1 and D_2 now may be regarded as special cases of the generalized dimension D_q for $q = 0, 1$ and 2. The spectrum D_q is a non-increasing function of q as shown in Fig. 2.3. Specifically, we have the following inequalities:

$$D_0 \geq D_1 \geq D_2. \qquad (2.9)$$

In general, when the spectrum D_q is q-invariant, a set is called monofractal. On the other hand, when the spectrum D_q is a decreasing function of q, a set is called multifractal. Specifically, if the spectrum D_q is constant for all $q < c$ and $q > c$ for some c, the set is called bifractal.
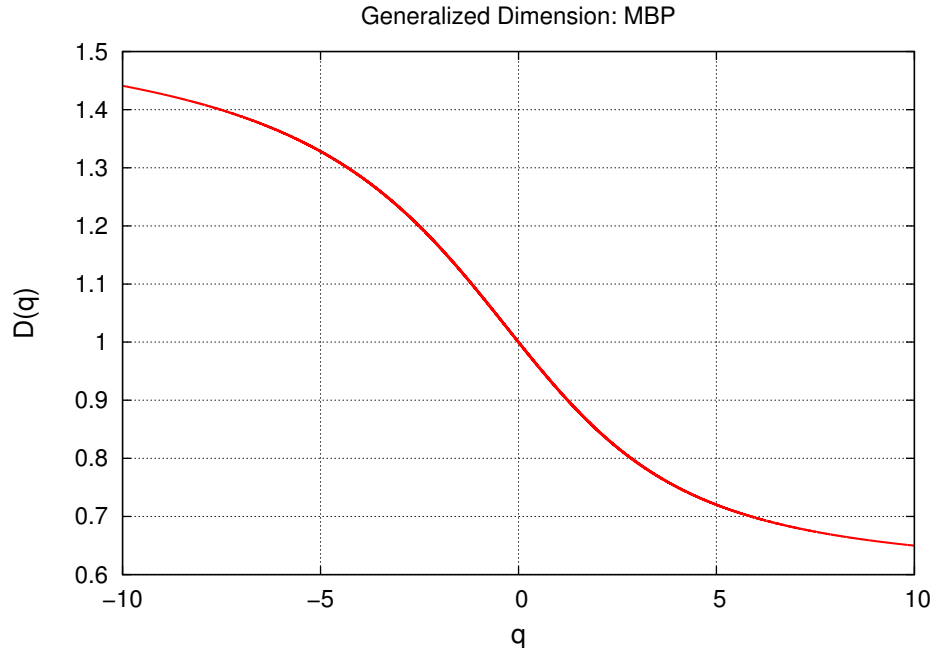
FIGURE 2.3: **The analytical result of the generalized dimension for the MBP.**

2.1.5 Multifractal

As briefly introduced in the last section, multifractal sets can be associated with different exponents depending on the power laws. A multifractal set arises when a mass distribution μ of a given set varies over the set and therefore is often studied in relation to dynamical systems. To see how a mass distribution contributes to the generalized dimension, consider a "mesh-like" cover B with each $|B_i| = \delta$. We assume there exists a spectrum function $f(\alpha)$ for each scaling exponent α such that for small ϵ, $N_\delta(\alpha)$, the number of δ-mesh cubes B_i with $\delta^{\alpha+\epsilon} \leq \mu(B_i) < \delta^\alpha$, is roughly of the order $\delta^{f(\alpha)}$ for small δ. In fractal literature, the scaling exponent α is also called the singularity and the associated function $f(\alpha)$ the singularity spectrum. Accordingly, the singularity function $f(\alpha)$ is given by [14]:

$$f(\alpha) = \lim_{\epsilon \to 0} \lim_{\delta \to 0} \frac{\log\left(N_\delta(\alpha + \epsilon) - N_\delta(\alpha - \epsilon)\right)}{-\log \delta}. \tag{2.10}$$

While it does not hold true in general, we can interpret the singularity spectrum as the box-counting dimension for a set such that

$$\left\{ x : \lim_{\delta \to 0} \frac{\log \mu\left(B(x)\right)}{-\log \delta} \to \alpha \right\}. \tag{2.11}$$

Note that Eq. (2.11) is the version found in most physics literature. This means that a multifractal set has an inhomogeneous mass distribution and a single exponent alone cannot fully characterize the complexity of the set. Heuristically speaking, we assign the singularity spectrum $f(\alpha)$ to a subset of the original set with different density associated with α.

It can be shown that the singularity spectrum $f(\alpha)$ is a convex function and is related to $\tau := (q-1)D_q$ by Legendre transformation [14]:

$$\tau(q) = q\alpha - f(\alpha) \tag{2.12}$$

$$q = \frac{d}{d\alpha} f(\alpha). \tag{2.13}$$

It means that if we know the singularity spectrum $f(\alpha)$, we can deduce the generalized dimension D_q or vice versa. The singularity spectrum $f(\alpha)$ for a multifractal set called MBP is shown in Fig. (2.4). The associated generalized dimension D_q is shown in Fig. (2.3).
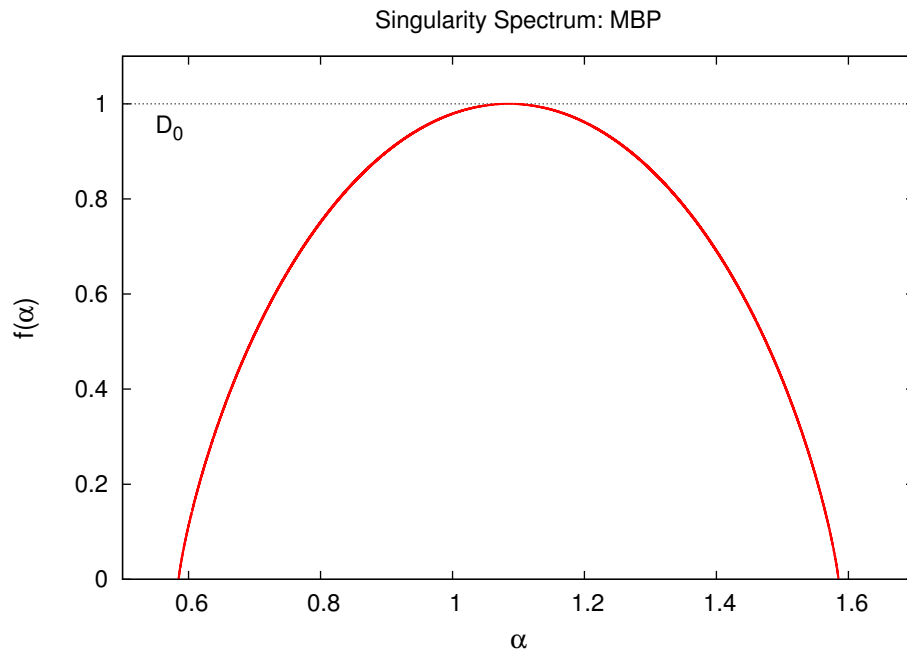
FIGURE 2.4: **The analytical result of the singularity spectrum for the MBP.**

2.2 The Cantor set

The Cantor set is one of the most iconic fractals and readily generalized to a multifractal set. Accordingly, we use the generalized Cantor set as our seminal test set to which the numerical methods are applied.

2.2.1 Definitions

The generalized Cantor set is constructed in the following way: starting with a interval of unit length, remove the middle part of the interval in such a way that the remaining interval on the left has a length of l_0 and on the right l_1. Moreover, a weight is assigned to each interval, namely p_0 or p_1 with $p_0 + p_1 = 1$ These weights determine the mass distribution of the set. The same procedure is applied to each of the two remaining intervals which then results in four intervals with lengths, starting from the left, l_0^2, $l_0 l_1$, $l_1 l_0$, l_1^2 and weights p_0^2, $p_0 p_1$, $p_1 p_2$, p_1^2. In general, after m such iterations, 2^m intervals with various factors are generated. A generalized Cantor set is what remains after taking $m \to \infty$. Particularly, a standard uniform Cantor set is obtained for $l_0 = l_1 = \frac{1}{3}, p_0 = p_1 = \frac{1}{2}$. Another special case,

referred to as the multiplicative binomial process, or MBP, is defined by $l_0 = l_1 = \frac{1}{2}$ with arbitrary weights [15]. Note that, unless $m = \infty$, the set is not a true Cantor set. For finite m, the set will be referred to as the finite representation of the Cantor set with hierarchy degree m.

2.2.2 Examples

Consider the generalized Cantor set with $l = l_0 = l_1$. Now, on the mth degree of hierarchy, the weight assigned to each interval is given by $p_k^{(m)} = p_0^{m-k} p_1^k$. The index k runs from 0 to m, depending on the location of the associated interval. Similarly, we can denote the length of each segment on the mth level by $l^{(m)} = l^m$. Then there exists the local dimension $\alpha_k \in \mathbb{R}$ such that $p_k^{(m)} = (l^m)^{\alpha_k}$. By extracting the exponent α_k from this relation, we obtain the following relation:

$$\alpha_k = \frac{\left(1 - \frac{k}{m}\right) \ln p_0 + \frac{k}{m} \ln p_1}{\ln l}. \tag{2.14}$$

The associated singularity spectrum $f(\alpha)$ is given by:

$$N_\alpha = \left(l^{(m)}\right)^{-f(\alpha)} \tag{2.15}$$

where N_α is the number of intervals having the local dimension α. Since the number of intervals N_{α_k} is simply

$$N_{\alpha_k} = \frac{m!}{k!(m-k)!}, \tag{2.16}$$

using Sterling's approximation, from Eq. (2.15) :

$$f(\alpha) = \frac{x \ln x + (1 - x) \ln(1 - x)}{\ln l} \tag{2.17}$$

where $x \equiv \lim_{m \to \infty} \frac{k}{m}$. After taking the limit for α_k in Eq. (2.14), we can obtain the implicit relation for $f(\alpha)$. Fig. 2.4 shows the spectrum for the MBP with $p_0 = 1/3, p_1 = 2/3, l = 1/2$. After the Legendre transformation, the generalized dimension D_q can also be obtained, and is shown in Fig. 2.3. By employing a similar argument, the spectrum as well as the generalized dimension can be found for a more general case, making the generalized Cantor set the ideal candidate for testing numerical methods.

Chapter 3

Numerical Methods

The box-counting method has been the most popular among researchers despite its difficulty to accurately compute the generalized dimension D_q [16] in the negative q range [17]. In this chapter, we revisit two known alternative numerical methods for obtaining generalized fractal dimensions and discuss their strengths and difficulties. Unlike the box-counting method [18], and the related correlation method [18], which employ partitions composed of equal-sized cells, the two methods examined in this thesis employ mass-oriented partitions. The nearest neighbor method [19] utilizes partitions composed of equal-mass cells while the k-neighbor method [20] uses partitions composed of cells with cumulative mass. These alternative approaches enable one to compute the generalized dimension on the domain where the box-counting method encounters difficulty. The treatment here closely follows our published work [12].

3.1   Numerical Methods

We introduce two mass-oriented methods for computing the generalized dimensions and the definition of Dimension Function which provides an alternative multifractal spectrum. The box-counting method is also briefly explained for comparison.

3.1.1 Box-counting method

This method is probably the most well-known and is closely related to the original definition of the Rényi Dimensions. The box-counting method is an umbrella term and so there are a few slightly different versions that fall under the name "box-counting methods", using "spheres" instead of "boxes" for example, but the underlying ideas are similar: generally, the number of cells required to cover the points in a given set, N, changes as the size of the partitions ϵ changes. By associating the mass distribution μ_i with each cell, the scaling relation can be procedurally computed from a fractal set as the size of the partitions decreases, namely,

$$D_q = -\frac{1}{1-q} \lim_{\epsilon \to 0} \frac{\ln \sum_{i=1}^{N(\epsilon)} \mu_i^q}{\ln \epsilon}. \tag{3.1}$$

Due to the simplicity of the method, it is widely used among researchers. However, it has been pointed out by many that this method and, more generally, methods that involve partitions of the same size such as the correlation method, do not work well for $q < 1$

3.1.2 The Nearest Neighbor Method

The approach called the "nearest neighbor method" was first introduced by Badii and Politi [19]. This method is essentially based on the assumption that there is an exponent D such that

$$<\delta> \sim n^{-\frac{1}{D}} \tag{3.2}$$

where $< \delta >$ denotes the mean distance from each point to its nearest neighbor among n randomly chosen points from a given test set. Comparing Eq. (3.2), with other scaling laws, the exponent D can be thought of as another definition of dimension. In fact, by naturally extending the premise, the Dimension Function $D(\gamma)$ can be computed by using the moments of order γ of the distribution function $P(\delta, n)$ generated by an ensemble of n randomly chosen points:

$$<\delta^\gamma> \equiv M_\gamma(n) \equiv \int_0^\infty \delta^\gamma P(\delta, n) d\delta = K n^{-\frac{\gamma}{D(\gamma)}} \tag{3.3}$$

where K is some function of n and γ which asymptotically remains bounded as n becomes large. Here, the meaning of γ should be clear; for larger values of γ the contribution of high-density regions is suppressed since they generate smaller values of δ, the distance to the nearest neighbor, and vice-versa. The proof of a more general relation is provided by van de Water and Schram [20]. Therefore, the nearest-neighbor approach may be regarded as a special case of more general scaling relations which will be discussed in Sec. 3.1.4. From Eq. (3.3), it follows that the Dimension Function $D(\gamma)$ can be obtained by:

$$D(\gamma) = - \lim_{n \to \infty} \frac{\gamma \ln n}{\ln M_\gamma(n)}. \tag{3.4}$$

The function K in Eq. (3.3) generally depends on n and γ but K should be, by definition, irrelevant in the limiting case as in Eq. (3.4). In numerical analysis, the value of $K(n, \gamma)$ does affect the numerical result since n is finite. The scaling property of Eq. (3.4) for the uniform Cantor set is shown in Fig. 3.1. The simulated results for $\gamma \ln(n)$ vs. $-\ln(M_\gamma(n))$ are plotted for a selected set of γ as n increases. For the 10 different values of n selected, the scaling property is clearly observed. In Fig. 3.2, the value of $D(\gamma)$ was extracted as the slope of the best-fit line in Fig. 3.1 for each corresponding γ. The slope values in the positive γ range agree well with the analytical result which is $\frac{\log 2}{\log 3}$ for the uniform Cantor Set [12].

3.1.3   Dimension Function

The Dimension Function $D(\gamma)$ can be thought of as an alternative generalized dimension and is related to the Rényi Dimension by [19]:

$$D[\gamma = (1 - q)D_q] = D_q. \tag{3.5}$$

As the equation suggests, once $D(\gamma)$ is obtained, the generalized dimension D_q can be found as the intersection of $D(\gamma)$ and the straight line with slope $(1 - q)^{-1}$ which passes through the origin as illustrated in Fig. 3.3. For most cases, the generalized dimension D_q is uniquely determined from $D(\gamma)$. Note that a larger q does not correspond to a larger γ due to the

The scaling of $\gamma\ln(n)$ vs. $-\log(M_\gamma(n))$ for the uniform Cantor set
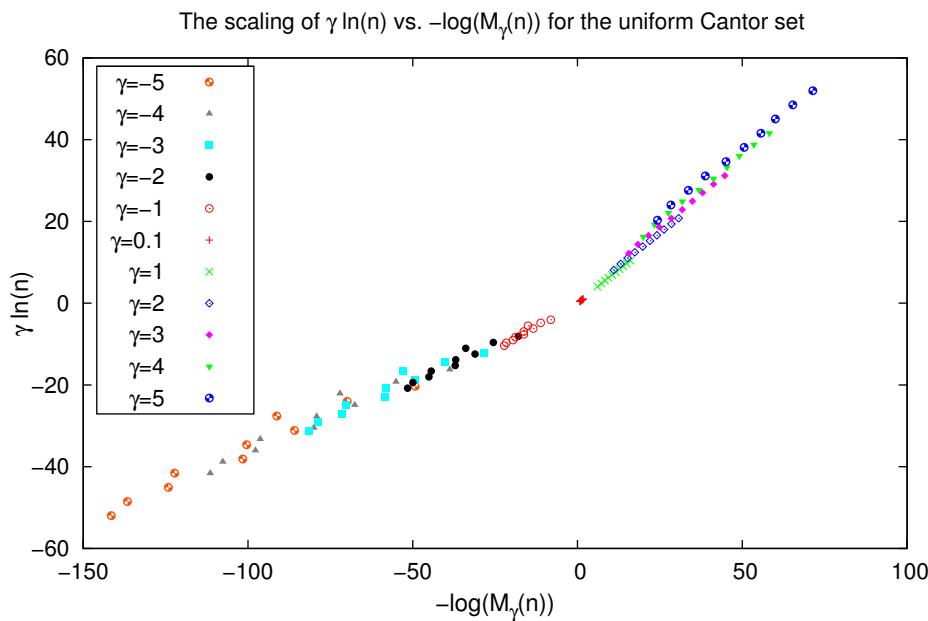
FIGURE 3.1: **For the uniform Cantor Set, $\gamma\ln(n)$ vs. $-\ln(M_\gamma(n))$ is plotted for each γ as n is increased. According to Eq. (3.4), the slope converges to $D(\gamma)$. The corresponding result for $D(\gamma)$ is shown in Fig. 3.2**

.

negative sign in the equation. However, the index γ still plays a similar role as q in that it discriminates the range of density of a given set that most strongly contributes to $D(\gamma)$.

In simulations, the Dimension Function $D(\gamma)$ is obtained using Eq. (3.4). The formula can, in principle, be applied to sets with any dimension of space a given set is embedded in. In the case of a one-dimensional set, sample points are prepared in a way that δ is bounded from above by 1. Therefore, the integral in Eq. (3.3) can be taken from 0 to 1. Unlike the box-counting method, this algorithm does not make use of partitions of the same size but, rather, of the same "mass" for it can be considered that each element of the partition contains two points, namely a reference point and its nearest neighbor. Badii and Politi used a slightly improved version of the method, namely "near-neighbor" method, which uses partitions containing three or four points to smooth out local statistical anomalies [19]. Broggi used partitions containing up to 300 points for systems of large dimensionality [21]. For all these approaches, the number of sample points in a cell is fixed while the total number of sample points n is increased when extracting the Dimension Function. Therefore,
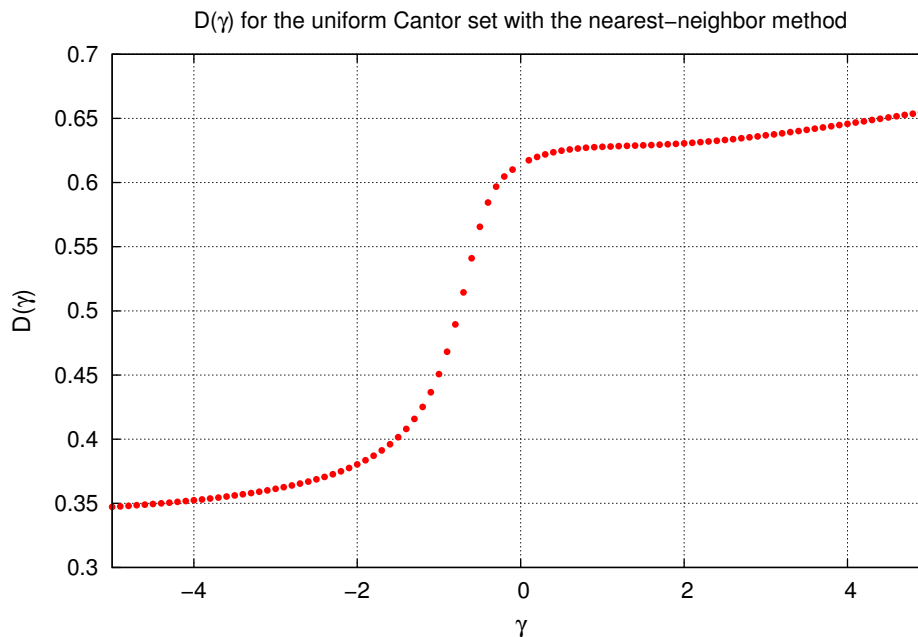
FIGURE 3.2: **In this graph, the Dimension Function $D(\gamma)$ for the Uniform Cantor set was computed as the slope of the best-fit line to the corresponding data set which is partially plotted in Fig. 3.1. $D(\gamma)$ diverges strongly from the analytical result which is $\log 2 / \log 3$ for negative γ.**

these methods differ from the k-neighbor method mainly in that the scaling of cell-size with n is used.

3.1.4   The k-neighbor Method

Another method, called "k-neighbor," is similar to the nearest neighbor method in that its partitions are taken according to the number of points inside. However, instead of fixing the number of points in a cell as in the case of the nearest neighbor method, the k-neighbor method incorporates a partition of cumulative points. Therefore, the scaling property is obtained through the global structure of a given set. A similar global approach with size-oriented partition was introduced by Tél *et al.* [22] using elements of different size, rather than different mass, and some literature misleadingly refers to it as the "cumulative mass" method [23].

The k-neighbor method records the distance $\delta(k,n)$ from a reference point to the kth neighbor point among $n-1$ randomly chosen points from a given set. van de Water and
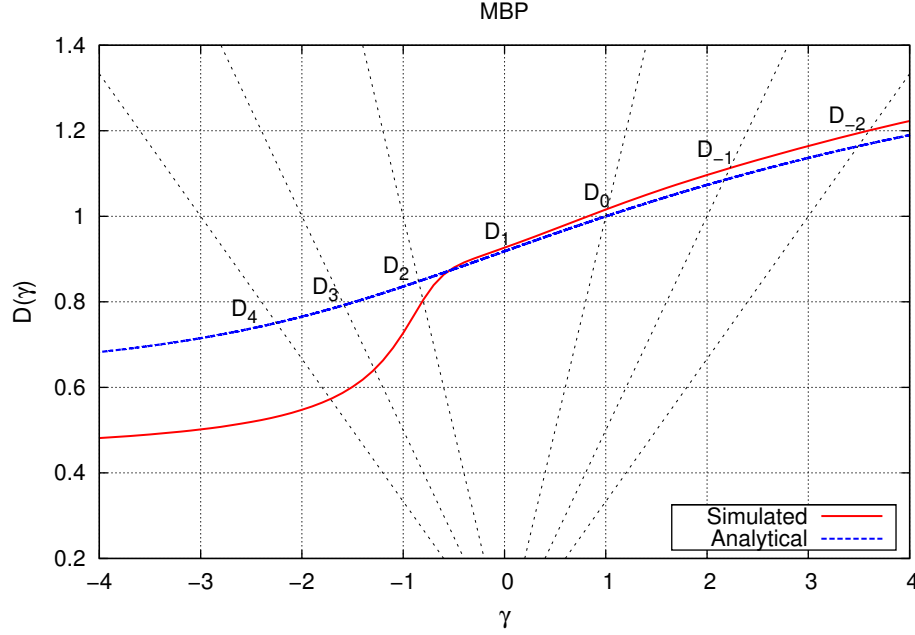
FIGURE 3.3: **The solid curve is the simulated result of the Dimension Function for MBP using the nearest neighbor method. Note how D_q can be obtained by locating the corresponding intersections. For example, the box-counting dimension D_0 can be found at the intersection of $D(\gamma)$ and $y = \gamma$.**

Schram formulated a technique for evaluating $D(\gamma)$ from the average of $\delta(k, n)^\gamma$ by using the local dimension introduced in Sec. 2.1.5 [20]. The average of $\delta(k, n)^\gamma$ is defined as follows:

$$\Delta^{(\gamma)}(k, n) = \frac{1}{n} \sum_{j=1}^{n} \delta_j^\gamma(k, n). \tag{3.6}$$

where $\delta_j(k, n)$ represents the kth neighbor distance from the jth reference point when n points are randomly chosen from a test set. Here, all n sample points are used as reference points. When n is large, it can be shown that [20]

$$\langle \Delta^\gamma(k, n) \rangle^{1/\gamma} \cong n^{-1/D(\gamma)} \left[\alpha D(\gamma) \frac{\Gamma(k + \gamma/D(\gamma))}{\Gamma(k)} \right]^{1/\gamma} \tag{3.7}$$

where α is some constant independent of γ. Note that the average of δ_j^γ from a single set is used in Eq. (3.6) whereas the derivation of Eq. (3.7) is based on the ensemble probability. For large k, a simple approximate relation can be obtained [20]:

$$\left[\Delta^{(\gamma)}(k, n) \right]^{1/\gamma} \cong n^{-1/D(\gamma)} k^{1/D(\gamma)} G(k, \gamma) \tag{3.8}$$
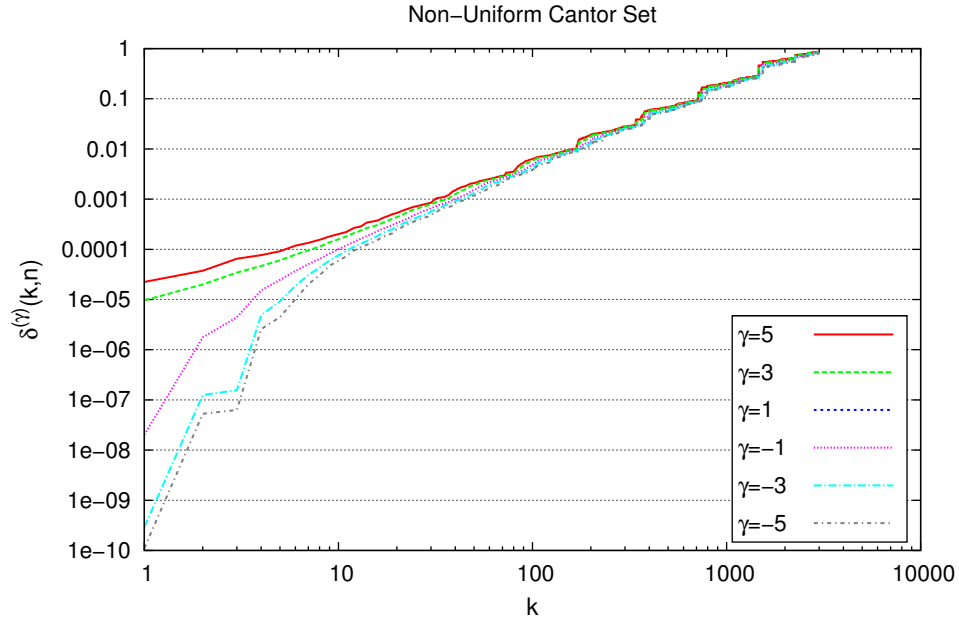
FIGURE 3.4: $\delta^{\gamma}(k,n) = (\Delta^{\gamma}(k,n))^{(1/\gamma)}$ **is plotted versus k in a log-log plot. The fine structure inherited from the non-uniform Cantor set is observed.**

where $G(k,\gamma)$ is a correction function close to unity for large k. The Dimension Function $D(\gamma)$ can be estimated by setting $G(k,\gamma) = 1$ in the first iteration. The dependence of the correction function $G(k,\gamma)$ on k and γ can be obtained from Eq. (3.7) with the value of $D(\gamma)$ from the first iteration. The Dimension Function $D(\gamma)$ then will be updated using this $G(k,\gamma)$. After a few iterations, the numerical results for $D(\gamma)$ will converge to a single value for each γ. The correction function $G(k,\gamma)$ generally exhibits a periodic pattern as a direct consequence of the self-similarity of fractals as seen in Fig. 3.4.

According to Eq. (3.8), the Dimension Function $D(\gamma)$ can, in principle, be obtained from the slope of the best-fit straight line in the log-log plot with either a fixed n or k. When k is fixed to 1, the equation reduces to the key relation in Eq. (3.3) for the nearest neighbor method. For the near-neighbor method, $k = 3$ or 4 may be used. With the k-neighbor method, we used a fixed value of n. By fixing n instead of k, we can extract a global property of a given set. The use of scaling with k makes the k-neighbor method less sensitive to local anomalies which often arise from a finite sampling process. It is worth noting that the Eq. (3.7) contains a gamma function which has singularities for $\gamma < -kD(\gamma)$. In general,

the bigger k we take, the larger is the valid range of the Dimension Function $D(\gamma)$.

3.2 Results

Generally, with a small amount of computational time, both of the methods in the fixed-mass class give good indications of the Rényi Dimension in the vicinity of the box-counting dimension ($q = 0$) on various generalized Cantor Sets [12]. This is a major advantage over the box-counting method if one seeks to find D_0, the box-counting dimension itself. Around the box-counting dimension, the nearest neighbor method yields a result closest to the analytical solutions. However, as q moves away from 0 and hence γ moves away from D_0, the k-neighbor method often produces more accurate results. Therefore, at this point, no single method seems reliable enough for an extended domain q of the generalized dimension. However, the combination of the aforementioned methods reveals the essential features of a given set such as whether it is a monofractal or multifractal.

For a multifractal set, how the dimension changes over the domain q is a key property. The k-neighbor seems to be the best method to start with as it can provide an estimate of the generalized dimension over an extended region, albeit not too accurately especially for $q < 1$. To obtain the dimension to a higher accuracy for a particular q or γ, the box-counting or the nearest neighbor method may be used. For $q > 1$, the box-counting method should be employed and for $q < 1$, the nearest neighbor, provided that q is not a very large negative number. Therefore, if possible, the results obtained from these methods should be compared and examined to see if they are consistent within the uncertainty of each method [12].

3.2.1 The Nearest Neighbor Method

In the nearest neighbor method, the Dimension Function $D(\gamma)$ was extracted from Eq. 3.4, where the right hand side reads $-\frac{\gamma \ln n}{\ln M_\gamma(n)}$ before taking the limit [12]. To investigate how it approaches to the limit, $\ln n / \ln M_1$ versus $\ln n$ for the uniform Cantor set was plotted in Fig. 3.5. The points in the plot indicate how $-\gamma \ln n / \ln M_1$ seemingly approaches the theoretical limit of $\ln 2 / \ln 3 = 0.63...$ as $\ln(n)$ increases in the case of uniform Cantor set.

However, it can be seen that the convergence rate is rather slow. Given that the hierarchy degree m is large enough, increasing n can almost always guarantee a higher accuracy around the box-counting dimension. However, since the convergence rate is rather slow, determining the limit is not a trivial task. For $\gamma = 1$, the number of sample points $n = 2^9 = 512$ was required to obtain the result within 5% accuracy and $n = 2^{17}$ to obtain the result within 3%. For quick simulations, we typically used $n = 2^{16}$ and 10 ensembles.
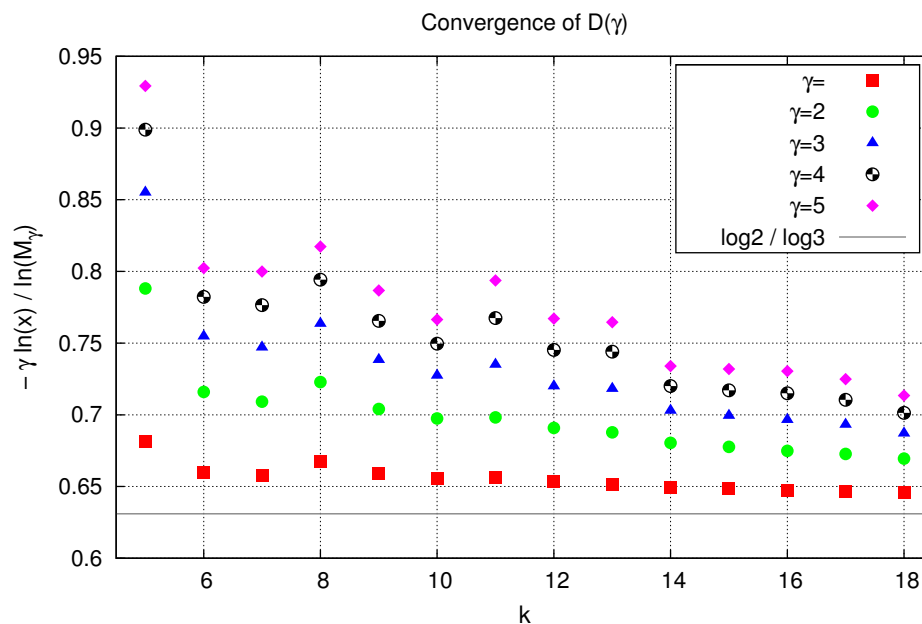
FIGURE 3.5: **This figure shows how increasing** $n = 2^r$ **affects the value of** $-\gamma \ln n / \ln M_\gamma$**. The plot was generated for the uniform Cantor set. The analytical value for** $D(\gamma)$ **for all** γ **is** $\log 2 / \log 3 = 0.630...$ **which corresponds to the horizontal line in the plot.**

In general, we employed the linear regression technique and obtained the limit from the slope of the appropriate log-log plot. While the overall qualitative features of the Dimension Function, such as the non-decreasing property, are properly reflected on the domain where γ is positive, the deviations and the fluctuations around $\gamma = -1$ seem sudden and uncontrolled. The difficulty of obtaining a sensible result for $\gamma < -1$ seems persistent throughout the set we have tested. In Fig. 3.6, the results for various generalized Cantor Sets are shown [12]; the domain of γ on which the simulated $D(\gamma)$ agrees well with the analytical results is between 0 and 2. For a multifractal, as γ increases, the numerical results start to diverge from the analytical result as well.
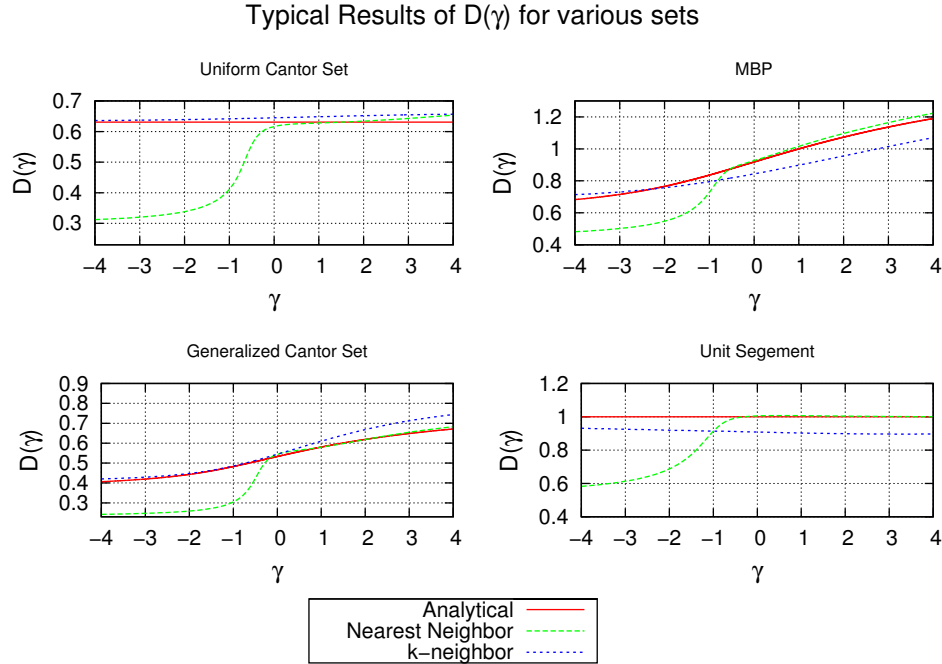
FIGURE 3.6: **These plots show typical results of $D(\gamma)$ vs. γ for the nearest neighbor method and the k-neighbor method applied to four different sets. The corresponding analytical results are shown for comparison. "Unit Segment" here means an interval of unit length and can be thought as the 0th finite representation of the Cantor set. For negative γ, numerical results persistently deviate from the analytical results for the nearest neighbor method. While the k-neighbor method works relatively well for all γ, the outcome may not be as accurate as the nearest neighbor method for small positive γ.**

3.2.2 The k-neighbor Method

Unlike the nearest neighbor method, where the choice of n is often limited by a finite sample size and the available computation time, the k-neighbor method can utilize a larger data set from which the slope is extracted to estimate $D(\gamma)$. In general fine structure occurs in the log-log plots which injects arbitrariness in a slope-fitting process. This point is covered in detail in Sec. 3.3.

For a fixed value of n, $D(\gamma)$ or, to be precise, the corresponding $1/D(\gamma)$ in Eq. (3.8), is taken as the slope of $\log \delta^\gamma(k, n)$ versus $\log k/n$ [12]. As clearly shown in Fig. 3.4, the $\delta^\gamma(k, n)$ obtained exhibits a periodic pattern, so all approaches for obtaining the slope seem to inject ambiguity. We have used the standard linear regression technique [24] using sample points equally spaced in the logarithmic scale of k rather than in the linear k scale. Another consideration is that the slope, and therefore the result for $D(\gamma)$, depends on the range over

which the linear regression is applied. It turns out that the best range seems to depend on the value of γ as shown in Fig. 3.7. The plot shows how $D(\gamma)$ varies when the upper bound of the slope range increases for the case of the uniform Cantor set with the analytical dimension of $\log 2/\log 3 = 0.63...$ for all γ values considered.
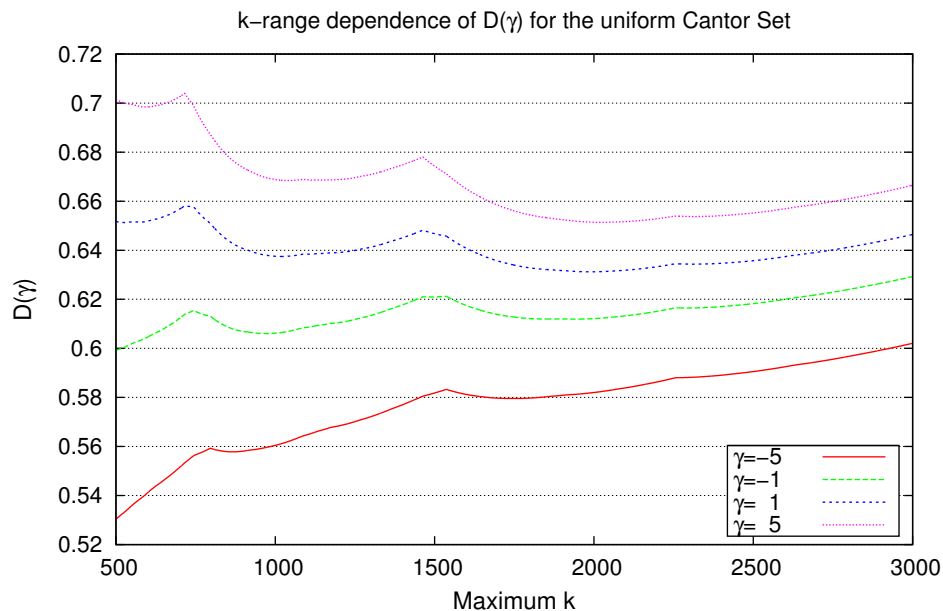
FIGURE 3.7: **This plot shows how $D(\gamma)$ differs when a different range is used to extract the slope in the k-neighbor method. For the uniform Cantor set, increasing the upper bound of k generally seems to produce better results. However, this is not a general result.**

As a result of these findings, we have used two different boundaries for computing the slope, one for positive γ and the other for negative γ, to produce the final results. Since the inaccuracy inherited from these ambiguities cannot be entirely removed by increasing n as in the nearest neighbor method, it is more difficult for the k-neighbor method to be adjusted to obtain a better result before knowing the theoretical values. Nevertheless, aside from these ambiguities in the method, the k-neighbor works for both positive and negative ranges of q, and therefore, is a good candidate as an initial method to investigate a given set. In the simulation, the ordering of the $n-1$ points for each reference point according to their relative position takes most of the computational time. Since the ordering takes more time as the embedding dimension increases, the method is said to be especially suited for one-dimensional sets.
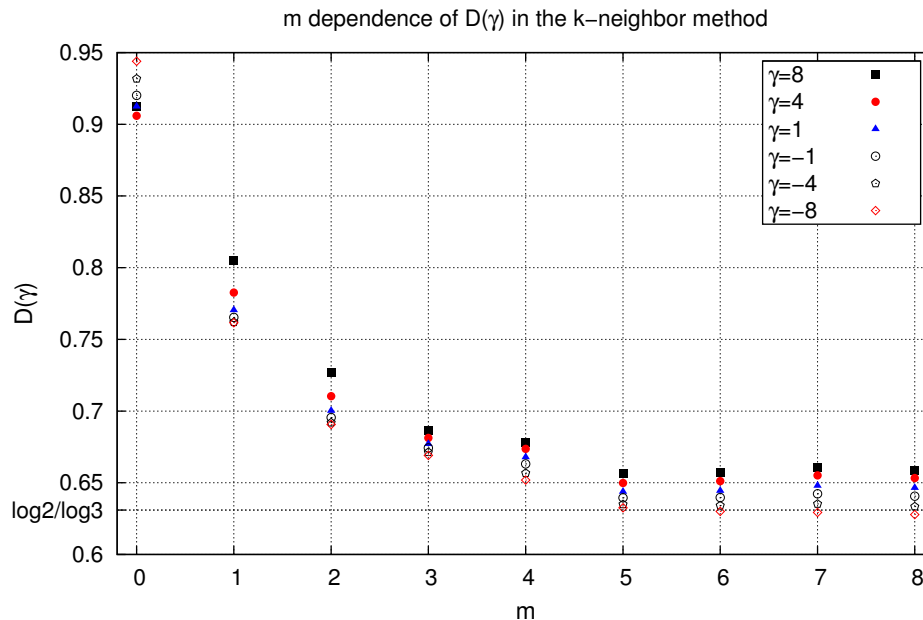
FIGURE 3.8: **These plots show how the results for $D(\gamma)$ change as m varies when the k-neighbor method is applied to the $m^{\mathbf{th}}$ finite representation of the uniform standard Cantor set. The theoretical value for $D(\gamma)$ is $\log(2)/\log(3)$ for all γ. For all iterations the value of n is fixed at 10000. The k-neighbor method provides relatively good results even when m is as small as 5.**

Furthermore, in contrast with the nearest neighbor method, the hierarchy degree m can be substantially smaller [12]. The size of the scaling region expectedly diminishes as m decreases. However, the Dimension Function deduced from the best-linear-fit from the appropriate scaling region produces acceptable results. For the uniform Cantor Set, when m is as small as 5, we obtained $D(\gamma)$ on the order of 0.6 as shown in Fig. 3.8. This shows that to estimate the fractal dimension from the k-neighbor method, the finite representation does not necessarily require a large degree of hierarchy. Hence, the k-neighbor method is a good candidate for estimating the fractal dimensions when only a limited hierarchy degree is available.

3.3   Analysis

3.3.1   Range and Stability

In the nearest neighbor method, the probability distribution of $P(\delta, n)$ plays a key role as seen in Eq. (3.3). Hence, it is worthwhile to investigate the nature of probability distributions associated with fractal sets. Starting with the conjecture for the mathematical form for the cumulative distribution function for the uniform Cantor Set,

$$S(\delta, n) = 1 - \exp[-n(2\delta)^{D_0}]. \tag{3.9}$$

Badii and Politi state that the correct form of the probability density distribution of uniform Cantor set for $n >> 1$ is given by [19]

$$P(\delta, n) = 2D_0 n(2\delta)^{D_0 - 1} \exp[-n(2\delta)^{D_0}]. \tag{3.10}$$

By substituting Eq. (3.10) into (3.3), a simple computation yields that

$$M_\gamma(n) = \left(\frac{1}{2n}\right)^{\gamma/D_0} \int_0^\infty x^{\frac{\gamma}{D_0}} e^{-x} dx \tag{3.11}$$

$$= \left(\frac{1}{2n}\right)^{\gamma/D_0} \Gamma(\gamma/D_0 + 1) \tag{3.12}$$

where $x = n(2\delta)^{D_0}$. Note that there is a singularity in the gamma function Eq. (3.13) for nonpositive integer z [25],

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt. \tag{3.13}$$

Therefore, the function $M_\gamma(n)$ involves singularities for $\gamma < -D_0$. This means that, for the generalized Cantor set, the nearest neighbor method is ill-suited for obtaining the Correlation Dimension ($q = 2$) or D_q for larger q. The function $D(\gamma)$ was calculated for each of four different data sets using the nearest neighbor method and is shown in Fig. 3.6. In each plot, the numerical results are compared to the corresponding analytical results. The influence of the singularity is observed for a variety of sets. Note that the k-neighbor method does not suffer from this kind of singularity. For the k-neighbor method, the corresponding

singularity can be found in Eq. (3.7). However, this time, the singularity can be avoided by taking a sufficiently large k. Accordingly, the k-neighbor method could generate sensible results in the entire range of γ we have investigated.

It is worth noting that the simulated probability distribution functions did not completely converge to the theoretical distribution of Eq. (3.10). The Komologov-Smirnov goodness-of-fit test [26], or simply K-S test, measures the maximum discrepancy between a sample cumulative distribution and a reference theoretical distribution. This test is often performed in order to determine if a sample distribution is drawn from the theoretical distribution. Here, the K-S test was employed to compare the theoretical distribution given by Eq. (3.9) with different values for D_0 and the distribution obtained in simulations [26]. As seen in Fig. 3.9, as m increases, the simulated distribution for the uniform Cantor set initially approaches the theoretical distribution when $D_0 = \frac{\ln 2}{\ln 3}$ is inserted in Eq. (3.9). When the number of points $n = 2^r$ exceeds the number of intervals 2^m, the nearest point for each reference point is likely to fall in the same interval which contains the reference point. This means that the nearest neighbor statistic does not reflect the property of the Cantor set but rather that of a line. Therefore, when m is increased, the result of the K-S goodness-of-fit test constantly decreases as long as $m < r$. One would rationally expect the convergence to improve when m is increased further but this was not observed. The maximum discrepancy reaches a plateau when $m = r$, suggesting that there is a constant disparity between the two distributions which does not diminish even when the finite representation of the Cantor set has a large hierarchy degree [12]. Among the values used, the theoretical distribution with $D = D_0 = \ln 2/\ln 3$ showed the best fit for $m > 14$. In the following, unless otherwise noted, we used the hierarchy degree of $m = 30$ when generating the finite representation of the Cantor set. Using a larger value does not significantly improve the results for the number of sample points we typically used, and would not comply with double precision arithmetic employed in our computations.

The effective domain is also related to the stability of the method. For both methods, as $|\gamma|$ increases, the nearest distance, δ, is either amplified or attenuated. Consequently, the contribution from only a few sample points among n chosen points starts to dominate the integral or sum in the equations. Unlike the nearest neighbor method, the effect of a few sample points is relatively small in the k-neighbor method due to the global feature. For the
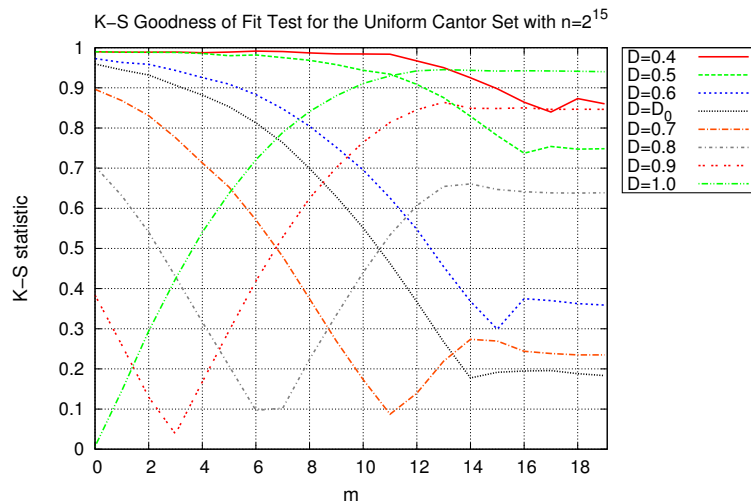
FIGURE 3.9: **The Kolmogorov-Smirnov goodness-of-fit test was used to compare the simulated probability density distribution and the theoretical distribution proposed by Badii and Politi for the uniform Cantor set with $n = 2^{15}$. According to Eq. (3.9), various values between 0 and 1 were substituted for D_0 for the purpose of this test. Smaller values of the outcome indicate a better fit. The finite representation of the Cantor set with $m = 1$ is the unit interval. Therefore, expectedly, the test function with $D = 1$ exhibits the best fit among others. As m increases, the K-S statistic decreases for $D = D_0 = \ln 2 / \ln 3$ and similar values. However, they reach plateaus after $m = 15$.**

nearest neighbor method, simulations require a large number of ensembles and therefore, an extensive amount of computational time and memory for a relatively large negative $|\gamma|$. How the Dimension Function $D(\gamma)$ varies in each implementation when the nearest neighbor method is applied is shown in Fig. 3.10 [12]. Each iteration is numbered on the horizontal axis. In the positive range of γ, the values of $D(\gamma)$ fluctuate more when computed under the same number of sample points as γ increases.

This difficulty can be partially overcome by employing the "near" neighbor instead of the nearest neighbor as it makes the simulation less dependent on the local property of a single reference point. However, it eventually suffers from the same difficulty as the magnitude of γ increases. The results for $D(\gamma)$ are shown in Fig. 3.11 when the near neighbor method is used [12]. The integer i denotes the ith neighbor points included in the partitions with $i = 1$ being the nearest neighbor method. Moreover, as i increases, all the relevant equations need to be modified accordingly but the dependence on i is not obvious. Overall, the k-neighbor method has an advantage for large $|\gamma|$.
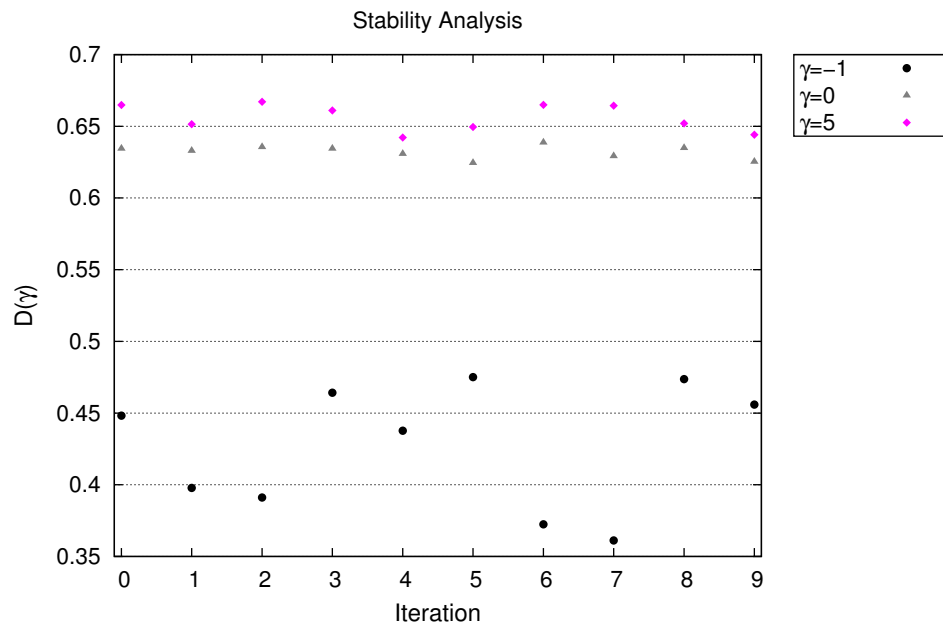
FIGURE 3.10: **This figure shows how each iteration of the simulation generates a different outcome for $D(\gamma)$. Each iteration is numbered on the horizontal axis. Sample sets were taken from the uniform Cantor set. In the range where Eq. (3.3) does not exhibit singularities, the results fluctuate more as γ increases. Larger fluctuation indicates more sensitive dependence on the particular choice of a sample set. The result for $\gamma = -1$ is also included to illustrate the difficulty of the method in the negative range of γ. The outcome in this range fluctuates even more and the average of the outcome is significantly smaller than the theoretical prediction which is roughly 0.63.**

3.3.2 The Limitation of Numerical Methods

As shown in Figs. 3.12 and 3.4, plots of the probability distribution $P(\delta, n)$ of δ for the nearest neighbor method or the kth neighbor distance $\delta^\gamma(k, n)$ typically exhibit self-similar fine structure which arises from the original fractal geometry [12]. However, unless a construction recipe is known in advance, as in the case of the generalized Cantor set, the exact nature of the fine structure is difficult to obtain. Moreover, to find its exact nature is essentially redundant for it would be another fractal set which is as complex as the original fractal set. Hence, numerical methods are typically developed based on an assumption that these fine structures will not affect their output in any substantial way. Nevertheless, we should not simply ignore the effect of the fine structure as a set would not be a fractal without them. In the equations such as Eq. (3.3) and Eq. (3.8), the fine structures are absorbed by
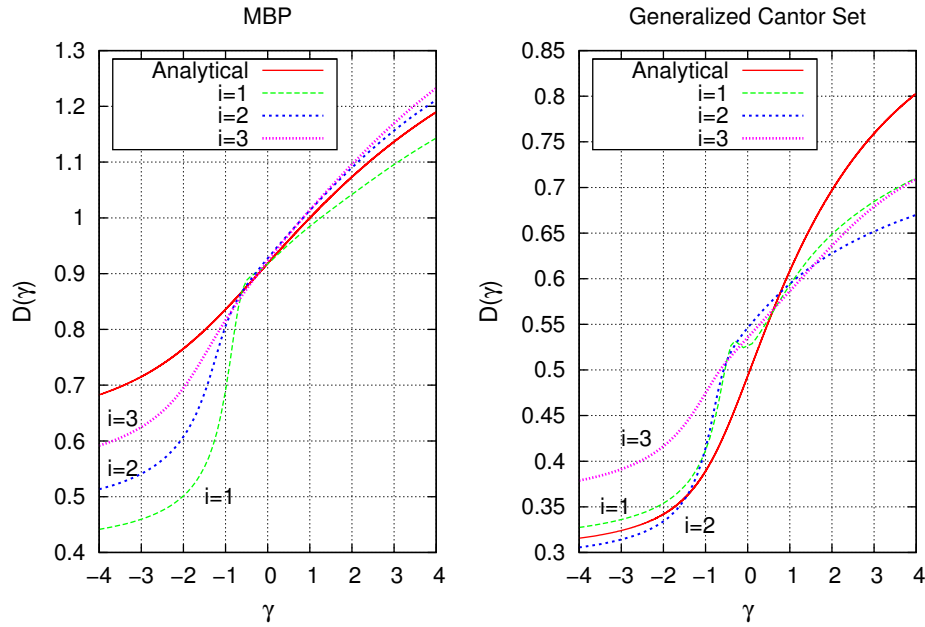
FIGURE 3.11: **These plots show how using near neighbor instead of the nearest neighbor affects the result. The integer i denotes the ith neighbor. While increasing i generally makes $D(\gamma)$ smoother, one cannot expect that the results improve when i is increased.**

the constant or correction term. In general, these correction terms depend on the hierarchy degree used in creating a test set as well as other parameters of these methods. However, it is difficult to estimate the error attributed to the correction term, and therefore this raises a question concerning the reliability of the method.

In principle, the largest possible m should be used to reflect the infinite hierarchical self-similarity. For the nearest neighbor method, the number of reference points, n, needs to be smaller than 2^m. Therefore, to increase n to obtain more accurate results, one needs to increase m as well. However, unlike the case of sample points where increasing n generally guarantees a more accurate result, increasing m does not necessarily do so. As we can see in Fig. 3.8, once m reaches a certain threshold, increasing m will not produce a better result [12].
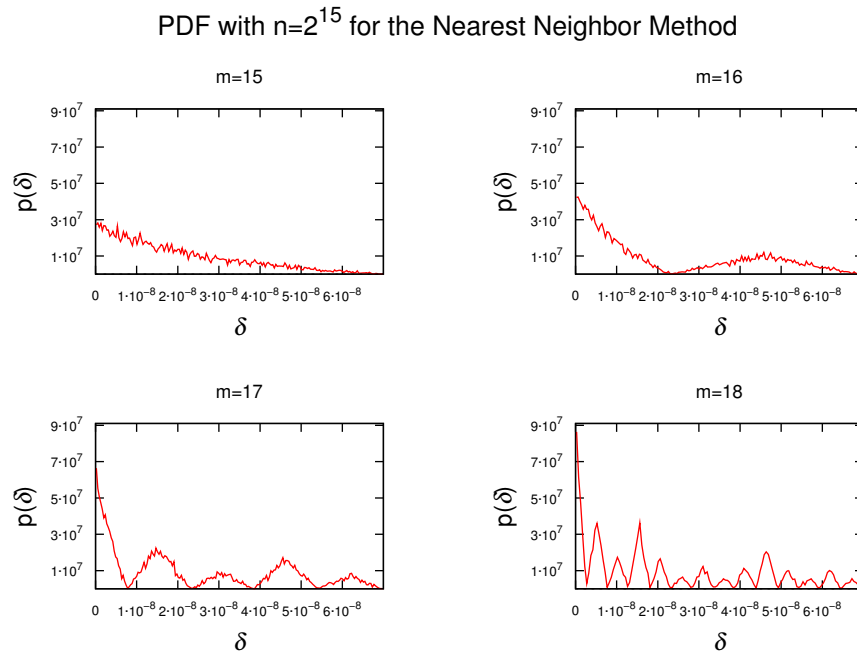
FIGURE 3.12: **These plots show how the hierarchy degree m affects the probability distribution of the nearest neighbor method. The sample sets were taken from the uniform Cantor set. While the cumulative distribution is somewhat more stable, as m increases, the fine structure of the probability distribution of δ emerges, exhibiting self-similar patterns. A limited horizontal range from 0 to 3^{-15} is plotted.**

3.4   Summary

In contrast with the box-counting method, or similar methods which utilize partitions into cells of equal size, the nearest neighbor method, which employs partitions of equal mass, as well as the k-neighbor method, which employs partitions of distributed mass, are good candidates for estimating the generalized fractal dimension for negative q. The k-neighbor method works for the complete range of q and no serious deviations were found [12]. By choosing an appropriate scaling region, we found that it is possible to estimate the generalized dimensions even with a small hierarchy degree [12]. However, the method involves linear regression and the results depend on how the best-fit line is obtained. Therefore, the k-neighbor method is a good option for a starting point and to investigate the general outlook of D_q.

If the sample size is large, the nearest neighbor method can be the best method for small negative q [12]. Although the result is sensitive to the local anomalies, one can choose the size of n according to one's required precision to extract the dimension. However, in contrast with the k-neighbor method, the hierarchy degree, m, also needs to be sufficiently large in order to obtain a desirable probability distribution. Therefore, if the sample size of a finite representation is small, the nearest neighbor method is not a practical choice. For positive q, the methods with partitions of equal size may be used.

In general, a few different methods should be applied before one determines if the results from different methods are consistent. The k-neighbor method should provide the overall features of D_q. Given that the subjective choice of the best-fit line affects the result, it is important to determine the window of ambiguity. If the sample size is adequate, apply the nearest-neighbor method for negative q and box-counting or similar method for positive q. The results from these two different methods should lie within the window of ambiguity. Given the strengths and the limitations of these methods, it would be interesting to apply them to a set with unknown fractal dimensions.

In any simulation of the kind worked out in this thesis, the finite sample correction needs to be taken care of. Although a number of correction terms have been proposed over the years [20, 27], many of them add extra complications to the simulation without achieving a dramatic increase in accuracy [20, 21, 28]. In the process of exploring the form of the nearest neighbor distribution of the generalized Cantor set, some interesting properties have been obtained; the order of taking m and n to infinity may not commute as usually assumed. Since a numerical sample only possesses a finite hierarchy, a new algorithm which does not assume an infinite hierarchy may be useful. In future work it will be shown that a new analysis of generalized dimension may be based on some quantities that are independent of the hierarchy.

Chapter 4

Cosmology

4.1 Introduction

In 1929, Edward Hubble discovered that the distant galaxies are receding from the earth approximately with a velocity v proportional to their distance [29]. If D is the proper distance to a galaxy, the Hubble's law may be stated as

$$v := \frac{dD}{dt} = H_0 D \tag{4.1}$$

where H_0 is the factor of proportionality called the Hubble constant and is responsible for the current expansion rate of the universe. This discovery led to wide acceptance of the expanding universe and replaced the idea of the stable universe. According to the standard theory of cosmology, the universe was hot and dense in its early phase [30]. As the temperature of the universe decreases, electronoically neutral atoms are formed, allowing photons to travel freely in space. The cosmic background radiation or CMB is believed to be leftover from this primordial universe. Current observation suggests that the CMB is nearly homogeneous and isotropic with deviations roughly 1 part in 10^5, confirming that the universe was homogeneous and isotropic in its early stage [31]. However, what we observe today in the universe appears to have large structures [32]. This observational trend indicates that as we probe further into space, we discover even larger structures.

Only recently, a large number particle simulations have become available to researchers [33] and therefore, structural formation in the universe is a highly researched area. The simulations show that a small fluctuation in the primordial universe can act as a seed and form a dense region in the matter-dominated epoch in a bottom-up manner [32]. This hierarchal structure formation spurs fractal analysis. As discussed in the previous section, fractal analysis requires a large number of test particles. In particular, we apply fractal analysis to a one-dimensional universe model. By confining the simulation to a single dimension, it is possible to significantly increase the number of particles per dimension to be computed. Our one-dimensional universe model simulation shows that the primordial universe with slight initial perturbations eventually creates large hierarchical structures as time progresses. In this section, we explain the formalism and assumptions we used to model our one-dimensional universe.

4.2   Friedmann Equation

The Friedmann Equation governs the expansion of the universe given the number densities of each species of substances. It can be derived from Einstein's field equation under some assumptions including the cosmological principle [34]. The equation reads

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G \rho}{3} \tag{4.2}$$

where a is the scale factor, k the intrinsic curvature parameter, G the universal gravitational constant, ρ the energy density. Roughly, the scale factor a represents the "size" of the universe. By eliminating the curvature parameter k, and assuming that the all energy is in one of the three forms, non-relativistic matter (M), radiation (R) and vacuum energy (Λ), we can rewrite Eq. (4.2) as [34]:

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0{}^2 \left(\Omega_M \hat{a}^{-3} + \Omega_R \hat{a}^{-4} + \Omega_\Lambda + (1 - \Omega_T)\hat{a}^{-2}\right) \tag{4.3}$$

where the density parameter Ω_i is determined by the ratio of the current value of the density to the critical density ρ_c ratio for each species i. Due to their properties, they

have characteristic dependencies on the scale factor. The parameter Ω_T in the last term is the current total mean density normalized to the critical density. Using the universal gravitational constant G as well as the Hubble's constant H_0, the critical density ρ_c is defined as [34]:

$$\rho_c := \frac{3H_0^2}{8\pi G}. \tag{4.4}$$

In the absence of the cosmological constant, the critical density represents a density threshold that determines the fate of the universe. The Friedman universe with a supercritical density will eventually start contracting and end with a "big crunch." On the other hand, a subcritical density means an ever expanding universe. In Eq. (4.3), the scale factor $\hat{a}$ is normalized to unity at the epoch of interest.

4.3 The Timeline of the Universe

Roughly, the timeline of the universe can be categorized by each epoch, radiation epoch; matter epoch: vacuum epoch; and curvature epoch. From the Friedmann equation, we can deduce the associated range of scale factor where each term dominates over other terms [35]. After the Big Bang, the universe first entered a radiation epoch, followed by a matter epoch. The future of the universe then depends on Ω_Λ as well as Ω_T. During the matter epoch, a small fluctuation in density can be the seed for a gravitationally bound structure. By neglecting the other terms, the dependence on t of the scale factor can be simply deduced. Therefore, from Eq. (4.3), the following relation holds [34]:

$$a(t) \propto t^{2/3}. \tag{4.5}$$

According to current theory, small fluctuations in density from the primordial universe act as the seed and eventually became what we observe today [32]. Our one-dimensional universe simulates this scenario to a certain degree.

4.4   Cosmological Principle

The cosmological Principle is that the universe is homogeneous and isotropic. The Friedmann equation is a solution to the Einstein Field equation using the Robertson metric which is based on the cosmological principle [35]. While the origin of this assumption is somewhat philosophical, various observations support this idea at least for large scale. However, as mentioned before, a small fluctuation that originated in the primordial universe grows in the matter dominated universe. The gravitationally bound structures are considered a local inhomogeneity. Some suggest that the universe may be fractal at all scales [36], indicating that the universe is scale-free at all scales. While a true mathematical fractal has an infinite nesting structure, and thus, is inhomegenous at all levels, most fractal-like structures found in nature have the lower and upper bound of the degree of hierarchy, outside of which may not be fractal-like. In fact, the standard ΛCDM theory suggests that the universe is homogenous at large scale. Isotropy may be violated in small scale as well. Dark matter is discussed in Sec. 4.5.

4.5   Dark Matter

Dark matter is a hypothesized particle which only weakly interacts with other particles via gravity. The existence of dark matter is deduced by its supposed gravitational effects on luminous matter which emits and absorbs light and therefore is visible to us [37]. More recently, the bullet cluster was found and is considered by some as a direct piece of evidence of dark matter [38]. The standard interpretation of the bullet cluster is that two clusters are undergoing a collision. The results with X-ray analysis and gravitational lensing show that the dark matter halo of each cluster is almost intact during the collision, passing through each other, while luminous matter are affected by the collision, losing energy. This observation further suggests that dark matter only interacts weakly.

The standard model of cosmology currently estimates that roughly 80% of matter in the universe is, in fact, dark matter [39]. Since what constitutes dark matter is yet to

be identified, this statement is still highly contentious. The estimate of the dark matter to luminous matter ratio is also constantly being updated by adjusting possible bias [40].

4.6   Scale-Invariant Universe

In a critical universe, a small over-dense region grows. To understand the evolution of matter distribution, consider the energy density $\rho(\mathbf{r})$ in a volume V of mean density $\langle\rho\rangle$. The density perturbation field $\delta(\mathbf{r})$ is then defined by [35]:

$$\delta(\mathbf{r}) \equiv \frac{\rho(\mathbf{r}) - \langle\rho\rangle}{\langle\rho\rangle}. \tag{4.6}$$

By expanding the perturbation field in a Fourier series, we have

$$\delta(\mathbf{r}, t) = V^{-1/2} \sum_{\mathbf{k}} \delta_{\mathbf{k}}(t) \exp(i\mathbf{k} \cdot \mathbf{r})) \tag{4.7}$$

where the sum is taken over all waves with $\langle\exp(i\mathbf{k} \cdot \mathbf{r}) = 0$. Note that the wave vector $\mathbf{k}$ is comoving in the expanding universe. The use of the density perturbation spectrum $\delta_{\mathbf{k}}$ allows us to separate large-scale effects, corresponding to small k, from small-scale effects, corresponding to large k. In an isotropic universe, the density perturbation spectrum must not have a preferred direction, giving a important scalar quantity called the power spectrum $P(k)$ [35]:

$$P(k) \equiv \langle|\delta_k|^2(k)\rangle = \langle|\delta_{\mathbf{k}}|^2(\mathbf{k})\rangle. \tag{4.8}$$

Perhaps, a more illuminating quantity might be the dimensionless function $\Delta(k)$ which is related to the power spectrum $P(k)$ in three dimensions by [35]:

$$\Delta^2(k) \equiv \frac{k^3 P(k)}{2\pi^2}. \tag{4.9}$$

In addition to the isotropy of the spectrum, if we assume that the spectrum does not possess a preferred length scale, the functional form of the power spectrum $P(k)$ is limited to a featureless power law [35]:

$$P(k) \propto k^n \tag{4.10}$$

with some index n. The conservation of momentum and asymptotic homogeneity require the index n to be [35]

$$-3 < n < 4. \tag{4.11}$$

Within the constrain Eq. (4.11), the value of $n = 1$ corresponds to the most important spectrum known as Harrison-Zel'dovich spectrum. To see the implication of this spectrum, in analogy to the energy contrast, expand the potential in a Fourier series [35]:

$$\Phi(\mathbf{r}, t) = V^{-1/2} \sum_{\mathbf{k}} \Phi_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r}). \tag{4.12}$$

The perturbation in potential $\delta\Phi$ is related to the density perturbation spectrum $\delta_{\mathbf{k}}$ via the Poisson equation,

$$\nabla^2 \delta\Phi = 4\pi G \rho_0 \delta. \tag{4.13}$$

Therefore, from Eqs. (4.8) and (4.13), we obtain:

$$\langle \Phi^2 \rangle = \frac{1}{2\pi^2} \int k^2 dk \langle \Phi_{\mathbf{k}} \rangle = \int_0^\infty k^2 dk \frac{P(k)}{k^4}. \tag{4.14}$$

It follows that the dimensionless quantity Δ_Φ^2 defined below is related to the power spectrum by

$$\Delta_\Phi^2(k, t) \equiv \frac{k^3 \langle |\Phi_{\mathbf{k}}|^2 \rangle}{2\pi^2} \propto \frac{P(k)}{k}. \tag{4.15}$$

Consequently, if $P(k) \propto k$, the potential is scale-invariant. This suggests that the corresponding metric is fractal as the potential perturbation dictates the flatness of spacetime.

If we assume that the CMB anisotropies caused by gravitational waves are negligible, the large-scale CMB anisotropies can be attributed to the inhomogeneity at the surface of the last scatterings of photons. The gravitational redshifts as a result of these scatterings are known as the Sachs-Wolfe effect. From the Wilkinson Microwave Anisotropy Probe (WMAP) via the Sachs-Wolfe effect, the potential fluctuation Δ_Φ in the cold-dark-matter-only universe is found to be [39]

$$\Delta_\Phi(k, t) = (2.97 \pm 0.06) \times 10^{-5}. \tag{4.16}$$

We introduce the spectral index n_s to account for the deviation of the index from the value

in Eq. (4.16), which indicates the tilt from the scale-invariance spectra. With correction functions $T(k,t)$ and $G(\Omega_M, \Omega_\Lambda, t)$ to the cold-dark-matter-only universe, the potential fluctuation may be written as [34]

$$\Delta_\Phi(k,t) = (2.97 \pm 0.06) \times 10^{-5} \left(\frac{ck}{H_0} \right)^{(n_s - 1)/2} T(k,t) G(\Omega_M, \Omega_\Lambda, t). \qquad (4.17)$$

The correction factors $T(k,t)$ and $G(\Omega_M, \Omega_\Lambda, t)$ are called the transfer function and the growth factor respectively. The transfer function T takes care of the suppressed amplitudes for large k during the radiation epoch. The growth factor is required when the density fluctuation stops growing when Ω_M is small. The value of n_s is estimated to be [39]

$$n_s = 0.960 \pm 0.013. \qquad (4.18)$$

The fact that the spectra index is close to unity means that the the primordial potential fluctuation is close to the scale-invariant Harrison-Zel'dovich spectrum. This is consistent with the inflation theory which predicts the primordial spectrum to be scale-free [41].

Chapter 5

Simulation

5.1 One-Dimensional Universe

In the past few decades, much numerical work simulating structure formation has been performed, most of which employs three-dimensional models. Though they have succeeded in reproducing the general features of the observed universe, these three-dimensional simulations are typically required to make a series of approximations which are not suitable for a fractal analysis. Here, we list several advantages of adopting a one-dimensional model instead.

First, most of the three-dimensional simulations employ Newtonian mechanics, thereby neglecting relativistic effects. By confining the particles in one dimension, we are no longer required to deal with the curvature of spacetime. Second, as we discuss later, the one-dimensional model allows us to construct a periodic boundary condition, which cannot be done in the same way using the Ewald summation with a three-dimensional model. Moreover, with three-dimensional models, due to the limitations of computational power, the number of particles per dimension is not good enough for fractal analysis. The most current work was carried out with a trillion particles, giving 10000 particles per dimension [42]. In this thesis, we used 10000 particles in our simulation, but one-dimensional models have a potential for computing larger number of particles. As mentioned in Chapter 3, fractal analysis typically requires a large number of particles, especially if we are required to analyze

the difference between each cluster in the system. Finally, in our one-dimensional model, the collision time between two particles can be solved analytically, making the algorithm event-driven. On the other hand, three-dimensional models typically resort to numerical solutions with a chosen time steps. Therefore, to obtain the results within a reasonable time, they often make approximations by inserting a cut-off in the potential and neglecting fine structures. These approximations may introduce unknown effects into the simulations. In addition, for fractal analysis, it is important to maintain the fine structures as they may play a crucial role in a fractal.

5.1.1 Formalism

In this simulation, we have adopted Newtonian mechanics with comoving coordinates. In the expanding universe, not only is the choice of the comoving coordinate theoretically the most natural choice, but also is computationally advantageous as we can work on the bounded region. While the same set of equations can be derived from the Friedmann equation, it may be worth including the Newtonian derivation here. In the expanding universe, the co-moving coordinate χ is often used in order to differentiate the motion of objets in the universe such as galaxy from the expansion and the peculiar motion. Introduce a scale factor $a(t)$ such that $r = a(t)\chi$. Then we have:

$$\frac{dr}{dt} = \chi \frac{da}{dt} + a(t) \frac{d\chi}{dt} \tag{5.1}$$

while the field equation gives

$$\frac{d^2 r}{dt^2} = E(r, t). \tag{5.2}$$

Therefore we have:

$$\frac{d}{dt} \left(\chi \frac{da}{dt} + a(t) \frac{d\chi}{dt} \right) = E(r, t) \tag{5.3}$$

$$2 \frac{d\chi}{dt} \frac{da}{dt} + \chi \frac{d^2 a}{dt^2} + a \frac{d^2 \chi}{dt^2} = E(r, t). \tag{5.4}$$

Since the gravitational field is a function of $1/r^2$, we obtain:

$$E(r,t) = (1/a^2)E(\chi,t). \tag{5.5}$$

Since we want to use an autonomous equation, for the matter dominated universe, we use the following scale factor:

$$a(t) = \left(\frac{t}{t_0}\right)^{2/3} \tag{5.6}$$

for some initial time t_0 which may be set to the beginning of the matter-dominated universe. That is, we have:

$$\frac{da}{dt} = \frac{2}{3}\left(\frac{t}{t_0}\right)^{-1/3}, \quad \frac{d^2a}{dt^2} = -\frac{2}{9}\left(\frac{t}{t_0}\right)^{-3/4}. \tag{5.7}$$

If we introduce a transformation $B(t)$ such that $dt = B(t)d\tau$ with a new time variable τ, then Eq. (5.4) becomes:

$$\frac{d^2\chi}{dt^2} + \frac{2}{a}\frac{d\chi}{dt}\frac{da}{dt} + \frac{1}{a}\chi\frac{d^2a}{dt^2} = \frac{1}{a^3}E(\chi,t). \tag{5.8}$$

To make the coefficient independent of time, introduce a transformation ξ such that

$$t = \xi(\tau), dt = \xi'(\tau)d\tau. \tag{5.9}$$

Here, the variable τ is an alternative time scale. Now, if we set

$$\frac{dt}{d\tau} = \frac{d\xi}{d\tau} = \xi' = \frac{t}{t_0}, \tag{5.10}$$

which yields

$$\frac{d^2\chi}{d\tau^2} + \frac{1}{3t_0}\frac{d\chi}{d\tau} - \frac{2}{9t_0^2}\chi = E(\chi,\tau). \tag{5.11}$$

Furthermore, let the unit of time be

$$1 = T_j \equiv \omega_j^{-1} = (4\pi G\rho)^{-1/2} = \frac{3}{\sqrt{2}}t_0 \tag{5.12}$$

where ω_j is the Jean's frequency. Note that Eq. (5.12) is a direct consequence of the matter-dominated Friedmann universe. Substituting this into Eq. (5.11), we obtain:

$$\frac{d^2\chi}{d\tau^2} + \frac{1}{\sqrt{2}}\frac{d\chi}{d\tau} - \chi = E(\chi,\tau). \tag{5.13}$$

This will be called the RF model after Rouet and Felix [43, 44].

5.1.2   Ewald Summation

To simplify the system, we embedded two-dimensional sheets with a mass density ρ that are aligned perpendicularly to the configuration space. Since they generate a constant field, we represent a sheet by a particle and focus on the position-independent potential along the one-dimensional space. However, with a periodic boundary $[-L, L)$, we also need to consider the potential from the replicas of each mass in the original range and therefore, it will diverge. To deal with this situation, we can use a technique called Ewald summation. We can show that the potential which contributes to the motion of the sheet is the remaining potential after the background potential is subtracted [45]. The background potential defined as the average potential over the periodic boundary.

Using the Ewald Sum, we can show that in the system with the number of parallel sheets $2N$ and the total length of the periodic system $2L$, the total field at a given position is given by

$$E(x) = 4\pi mG \left[\frac{N}{L}(x - x_c) + \frac{1}{2}\left(N_R(x) - N_L(x)\right) \right] \tag{5.14}$$

where x_c is the center of mass of the system and $N_{R(L)}$ is the number of particles to the right (left) of that position on the segment $[-L, L]$.

5.2   Algorithm

In our system, each particle has the same mass. The initial positions are chosen using the random number generator by sampling the theoretical distribution. Since this is a one-dimensional system, for a given particle, an "encounter" can occur only with neighboring

particles. Accordingly, we compute the encounter time for a given particle with neighboring particles. Solving for encounter times amounts to solving cubic equations. Therefore, the encounter time can be made precise almost up to the long double precision. We then single out the particle with the earliest encounter time. In our system, an encounter between two particles is simulated by passing of two particles. An encounter or "collision" involving two luminous particles results in the loss of energy after passing. In other case, the velocities of two particles are chosen in a way that the momentum as well as energy of the two particles are conserved.

To determine the energy loss between colliding luminous particles, we implement the restitution coefficient. If the difference in the velocities of the two passing particles is relatively large, the energy loss is also large, while the difference is relatively small, the energy loss is also small. This is qualitatively consistent with the standard theory [46]. At the encounter with the earliest encounter time, the new velocities are assigned to each particle. Since the ordering of the particles is swapped, new encounter times need to be computed for the encountered particles as well as the two neighboring particles. Since our system is periodic, the leftmost particle in a given interval may interact with the rightmost particle. In other words, the leftmost (rightmost) particle may exit the interval from the left (right) and re-enters from the right (left).

5.3   Numerical Challenges

In this one-dimensional model, the ordering of the particles is crucial. Therefore, numerical challenges arise when the distances between particles within a cluster become so small that the ordering is not properly preserved. This issue accelerates even more when luminous matter is introduced into the system as it tends to lose energy relatively faster. A closely related numerical challenge includes an encounter involving more than two particles. The current program can handle an encounter between two particles only. When particles are located close together, encounters involving multiple particles take place more often. After two particles collide, new velocities and positions are assigned to them. Analytically, this process should not destroy the ordering of the particles in the system as the positions of two

particles immediately before and after the encounter are identical. In the numerical simula-
tion, however, they may not be exactly the same, opening up the possibility of mishandling
the ordering. For instance, when a third particle collides almost at the same time, the new
position assignment to the two particles may alter the relative position between the two
collided particle and the third particle. Consequently, this third particle becomes a "ghost"
particle and eventually escapes from the system. Inevitably, the program crashes when too
many multi-particle-encounters take place.

Furthermore, depending on the scheme in which luminous matter loses its energy
in interaction with itself, an inelastic collapse may occur. While inelastic collapse can be
avoided by implementing a velocity dependent restitution coefficient, the average number
of encounter per unit time exponentially increases. This makes it difficult to follow the
evolution of the system after some t, within a reasonable amount of CPU time. In general,
the computation takes exponentially longer as the system evolves as shown in Fig. 5.1.

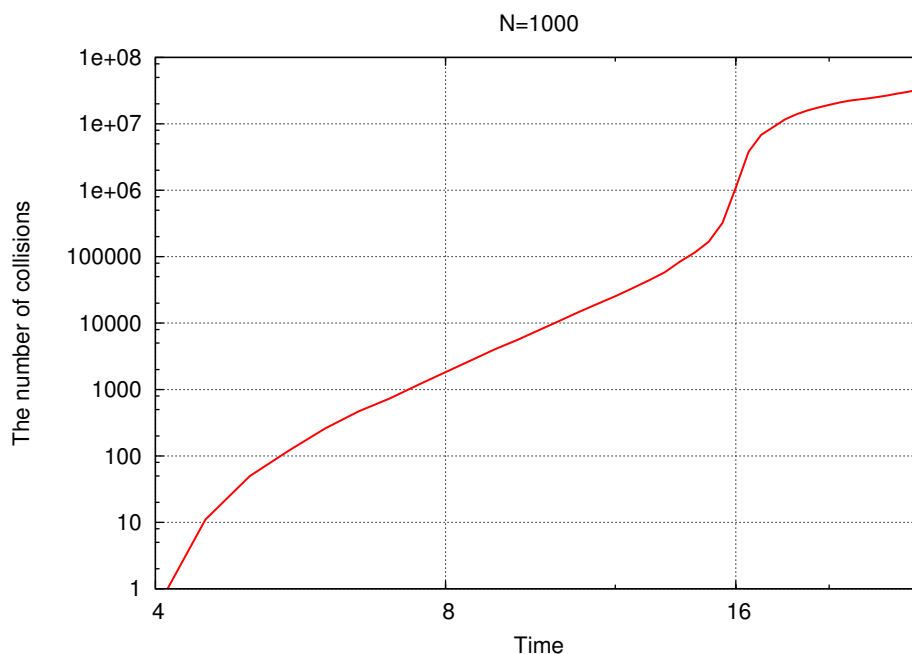FIGURE 5.1: **The number of encounters at a given time for a particular instance
of the simulation with $N = 1000$ is shown in a log-log plot.**

One way to deal with these difficulties is to treat a very dense cluster as one
particle with multiple mass. This will significantly reduce the computational time and
multiple collisions are less likely to happen. By adopting this method, we lose some pieces

of information within the very dense clusters, and therefore, the super-fine structure of the fractal distribution. However, unlike the numerical approximations often employed in the three-dimensional simulations, it will not generally affect the rest of the system. Moreover, since we can simulate longer time periods of the evolution, we can allow the rest of the system to fully grow and obtain better statistics of dense regions of the simulated set overall.

5.4 Initial Parameters

As explained in Chapter 4, in the primordial universe, it is believed that there were small fluctuations in energy density ρ. The variance of primordial density contrast δ is typically quantified by a power spectrum $P(k) \equiv <|\delta_k|^2>$. The one-dimensional analogue of Eq. (4.14) is given by

$$\langle \Phi^2 \rangle = \int dk \langle \Phi_{\mathbf{k}} \rangle = \int_0^\infty dk \frac{P(k)}{k^4}. \tag{5.15}$$

In this case, the dimensionless quantity Δ_Φ^2 is related to the power spectrum by

$$\Delta_\phi^2(k, t) \equiv k\langle |\Phi_{\mathbf{k}}|^2 \rangle \propto \frac{P(k)}{k^3}. \tag{5.16}$$

Therefore, in a one-dimensional universe, the power spectrum

$$P(k) \propto k^n \tag{5.17}$$

with n=3 corresponding to the scale invariance spectrum. In this work, the simulation was run with the fixed value of $n = 3$. We randomly assign the initial positions of the mass sheets according to the initial distribution function

$$x_j - x_{j0} = C \sum_{k>0} k^{n/2-1} \left[\sin(kx_{j0} + \theta_k) - \sin\theta \right] \tag{5.18}$$

for $k = \nu\pi/N$ with $\nu = 1, \ldots, N$.

5.5 Equation

The equation of motion for the jth particle is given by solving the following.

$$\frac{d\chi_j}{d\tau} \equiv v_j; \frac{dv_j}{d\tau} + \frac{1}{\sqrt{2}}v_j = E(\chi, \tau) \tag{5.19}$$

The phase space of a typical early state of the system is plotted in Fig. 5.2. The blue dots represent dark matter sheets at the position and red dots luminous matter sheets. In this plot, the number of total particles was chosen to be $N = 10000$. At a later time $T = 20$, as seen in the Fig. 5.3, the dots start to coalesce together in the phase diagram, with somewhat vertically elongated distributions. This means that there are some variations in velocities within the clusters.

FIGURE 5.2: **The distribution of matter particles at an early stage ($T = 0.5$) is plotted in phase space. The blue dots represent dark matter, red dots luminous matter.**

In Fig. 5.4, a zoom of the shaded area indicated in Fig. 5.3 is shown. Here, we see that, each cluster typically contains both dark matter and luminous matter. Surprisingly, some of the clusters appear to consist entirely of luminous matter. This suggest that the
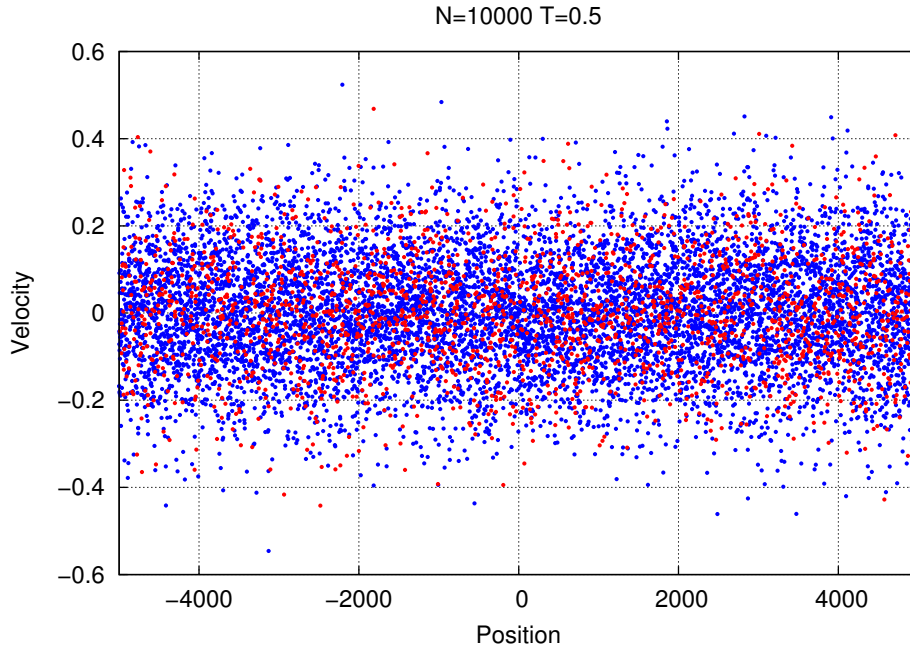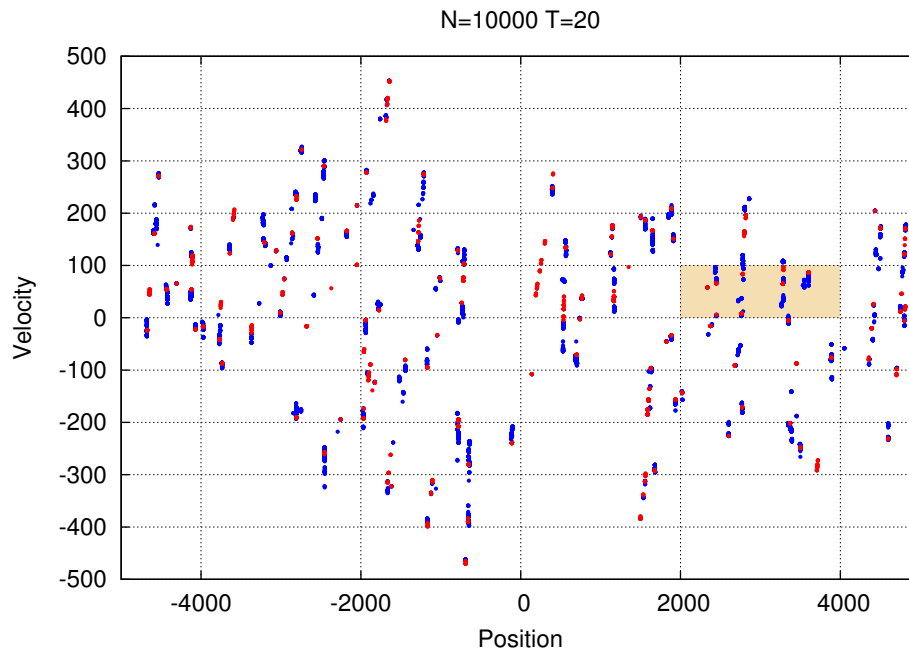
FIGURE 5.3: **The distribution of matter particles at a later time ($T = 20$) is plotted in phase space. The blue dots represent dark matter, red dots luminous matter.**

resolution of the plot may be beyond what the simulation can handle at this time. Fig. 5.5 shows the further zoom of the shaded are indicated in Fig. 5.4. Along with Fig. 5.6, we see that each cluster has rich fine structures within it. While we cannot tell if the distribution is fractal or not by examining these phase space plots alone, magnified plots reveal the hierarchical structure being formed.

Due to the loss of energy, luminous matter tends to coalesce together faster, resulting in its high concentration at the center of clusters. This can be clearly seen in the corresponding density distribution in Fig. 5.7. As T progress, we observe a few clusters in the system. As shown in in Fig. 5.6, the clusters start coming together, with a few stray clusters still present. Although it is not clear from this plot alone, each cluster appears to form a fractal set. The probability distribution of each matter species shows that luminous matter occupies the core of clusters while dark matter is more spread out around the core. The early distribution is almost homogenous as shown in Fig. 5.8. At $T = 15$, the distribution becomes less homogenous as seen in Fig. 5.9. At $T = 20$, several clusters are formed. As seen in Fig. 5.7, luminous matter has relatively a large peak at the core of the

clusters. By comparing it to the total distribution, at the core of the clusters, luminous matter dominates although there are more dark matter than luminous matter.
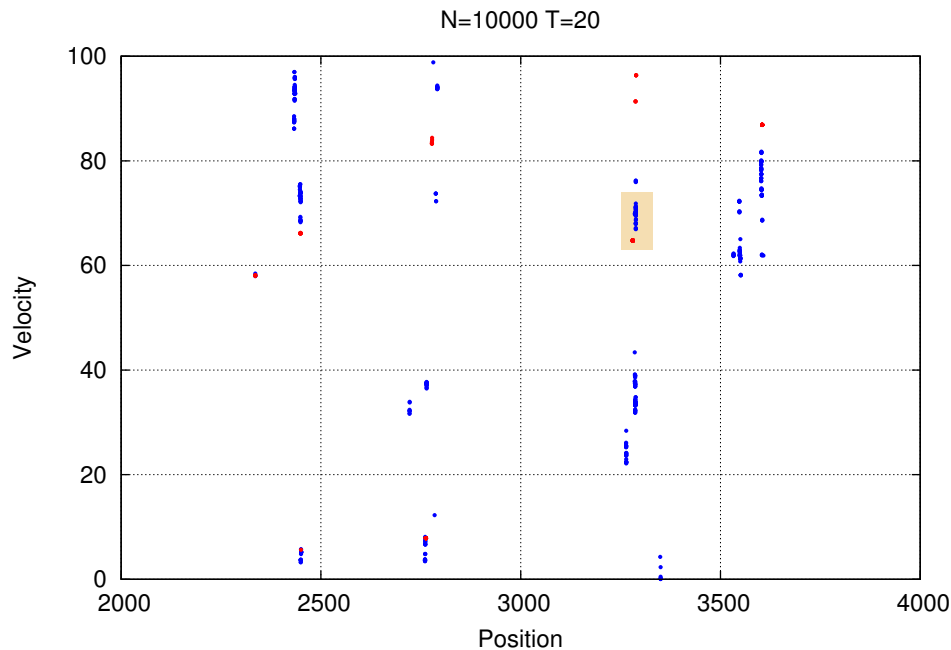

N=10000 T=20

FIGURE 5.4: **A zoom of the shaded area in Fig. 5.3 is shown. We observe a pattern similar to Fig. 5.3, suggesting that the set is a fractal.**

5.6   Preliminary Results

In this section, we present preliminary results of fractal analysis. We applied the fractal analysis methods mentioned earlier to the result of the simulation.

The typical result from the k-neighbor method for $T = 20$ shows that the distribution of matter is highly fractal and is multifractal as seen in Fig. 5.10 . Even though the result is not always quantitatively identical due to the difficulty of choosing an objective correct scaling range, the general trend is clear: for small T, the fractal analysis shows that the set resembles that of a line segment, meaning that the fractal dimension is one and uniform. As T progresses, multifractality evolves. Generally, at one point, all the lines with different γ start to converge and the dimension associated with this large k range is close to 1. This suggests the bottom-up evolution of the structure and that at large scale, the universe is homogeneous. This homogenous point increases in k as the system evolves.
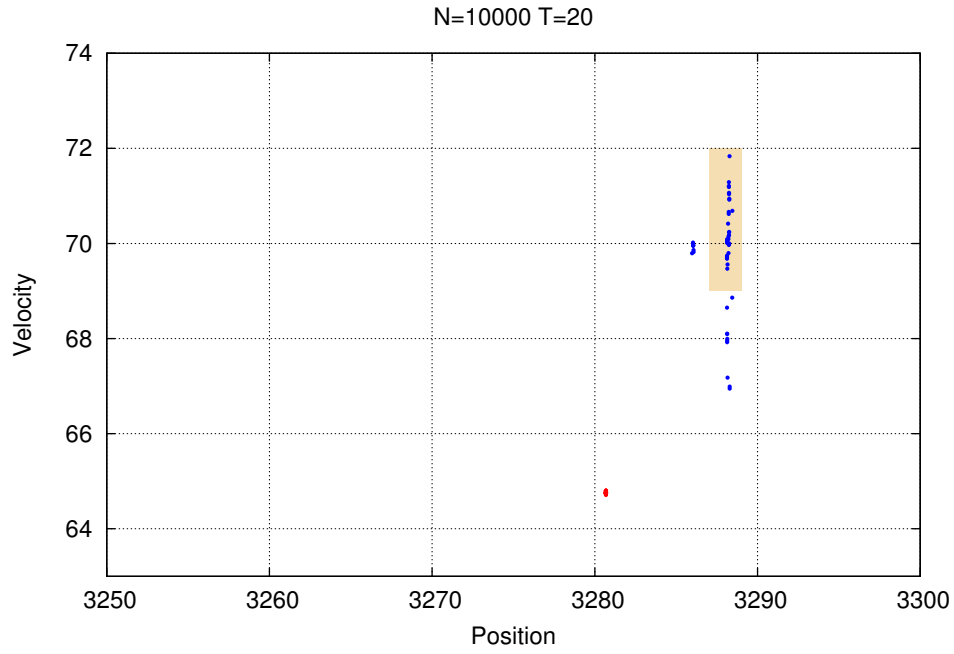
FIGURE 5.5: **A zoom of the shaded area in Fig. 5.4 is shown. We see two clusters consisting only of dark matter and of luminous matter.**
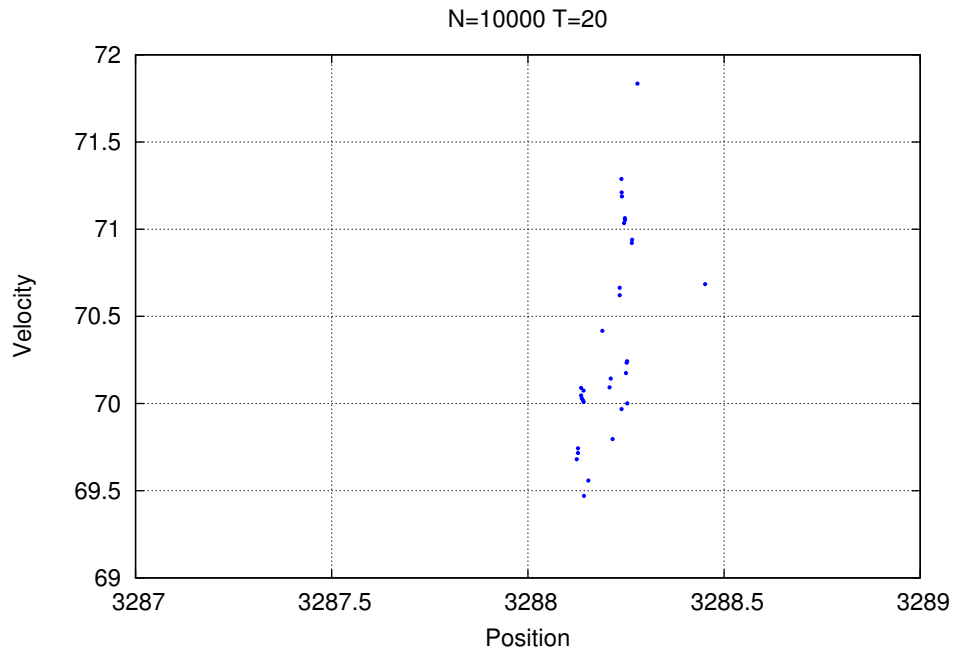
FIGURE 5.6: **A zoom of the shaded area in Fig. 5.4.**

Probability Distirbution (N=10000)
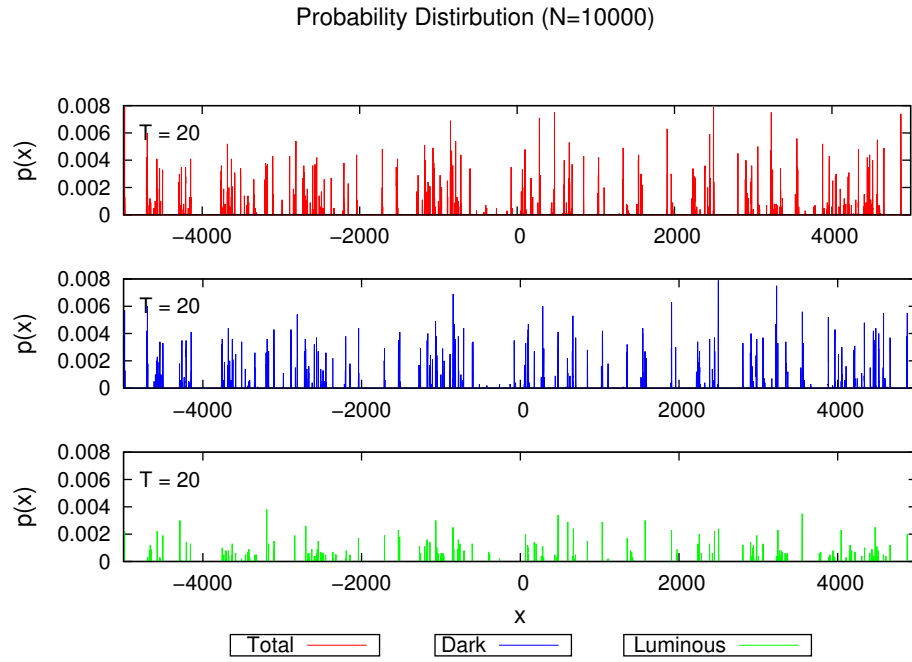
FIGURE 5.7: **The probability distribution of both matter as well as each species at** $T = 20$**.**

Probability Distirbution (N=10000)
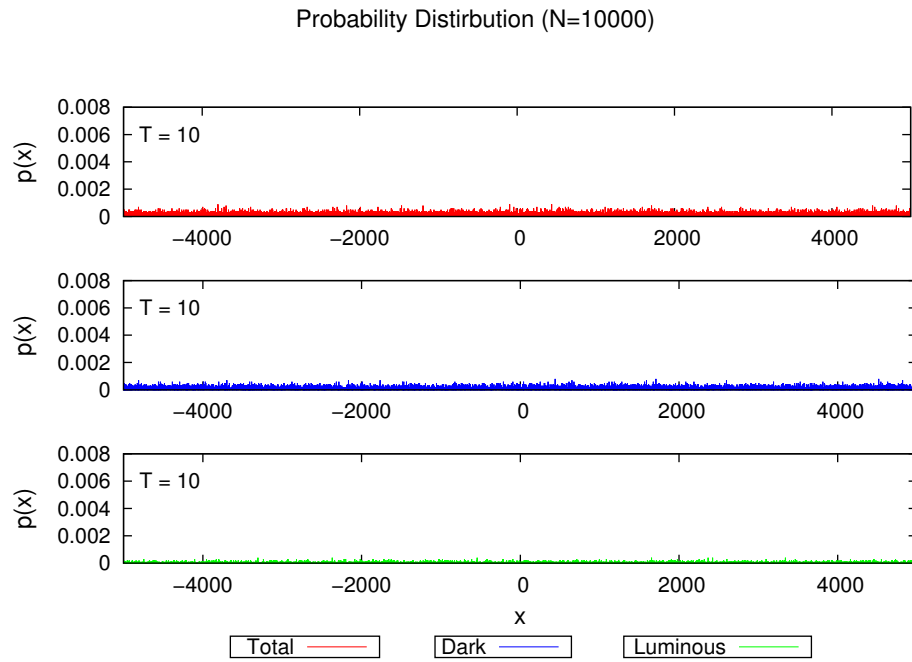
FIGURE 5.8: **The probability distribution of both matter as well as each species at** $T = 10$**.**

Probability Distirbution (N=10000)
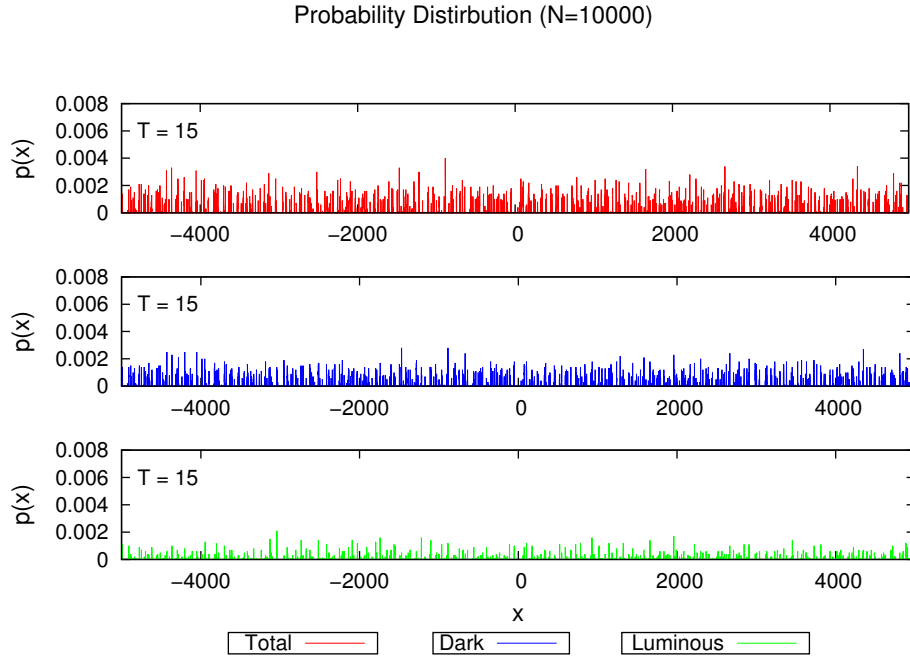
FIGURE 5.9: **The probability distribution of both matter as well as each species at** $T = 15$**.**

While the location of this homogenous point is rather clear from the graphs, the choice of a lower boundary for the scaling range is somewhat subjective. As in the case of the Cantor set, the scaling plots exhibit oscillatory patterns which are the hallmark of fractal sets.

In Fig. 5.11, we plotted the scaling of $\delta^\gamma(k, n)$ in a log-log plot. Here, we can observe several possible scaling regions. We can see that around at $k = 10$, the lines with different values of γ merge. For $k > 10$, the Dimension Function $D(\gamma)$ is almost uniform and has the value close to unity, suggesting that the distribution becomes homogenous for $k > 10$. This value of $k = 10$ appears to be independent of the number of particles in the system. With fewer particles, the point of merger is consistent at a given time. As T grows, the point above which the distribution is homogenous increases. When the system reaches roughly $T = 20$, the Dimension Function $D(\gamma)$ significantly drops in the negative γ region, as shown in Fig. 5.12.

The difficulty with extracting the fractal dimension with the k neighbor method is evident. It is not clear where the scaling regions are. In general, we can roughly divide the graph into three regions. Since the scaling range for positive values of γ seems to cover the whole range, the three divisions are made according to the behaviors of $\delta^{(\gamma)}$ with

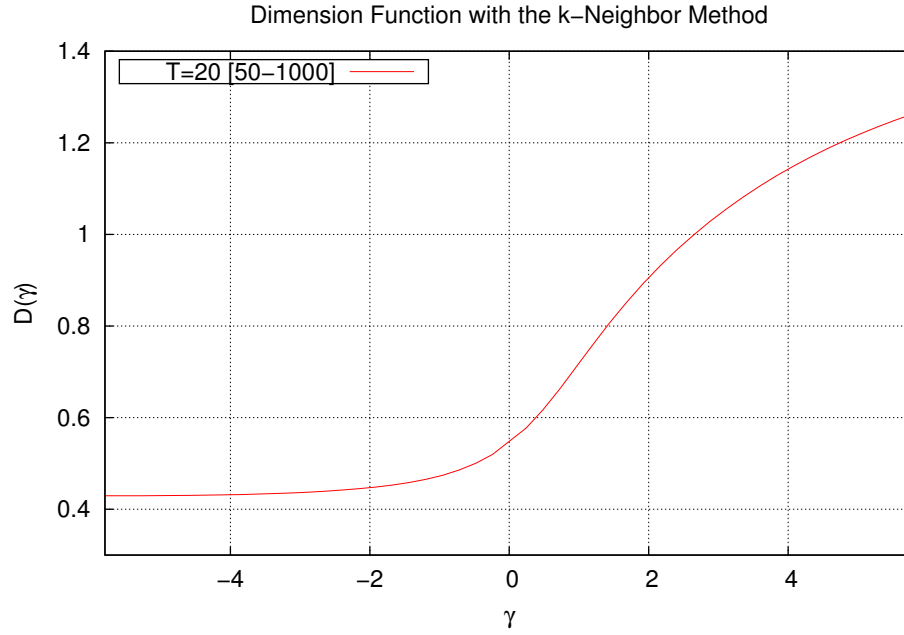Dimension Function with the k–Neighbor Method

FIGURE 5.10: **The Dimension Function $D(\gamma)$ was computed from the k-range between 50 and 1000. The corresponding scaling plot is shown in Fig. 5.12.**
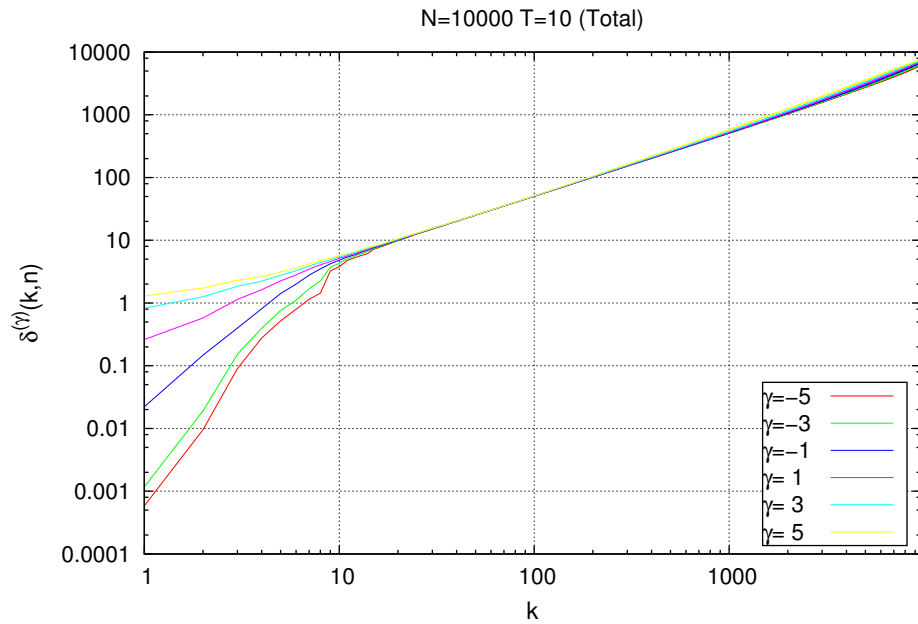
N=10000 T=10 (Total)

FIGURE 5.11: **The scaling of the kth neighbor distribution with both kinds of matter at $T = 10$.**
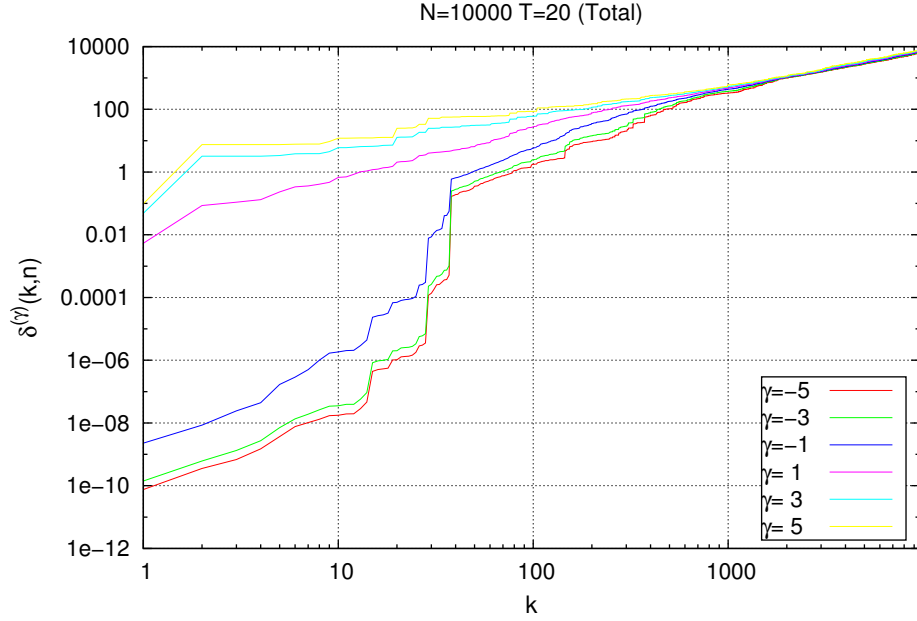
FIGURE 5.12: **The scaling of the kth neighbor distribution with both kinds of matter at $T = 20$.**

negative values of γ. In this particular case with $N = 10000, T = 20$, the first region ends at about $k = 40$. The spectrum of dimension was extracted from the second region. It starts approximately from $k = 40$ and ends at $k = 600$. For $k > 600$, all the lines appear to converge in the log-log plot.

While the kth neighbor distribution is shown in Fig. 5.12 when we do not discriminate the two species of matter, a closer look reveals that the two kinds of matter have distinctive distributions. The dark matter appears to exhibit a typical multifractal spectrum as shown in Fig. 5.13. On the other hand, luminous matter shows quite different behaviors between positive and negative values of γ. While the Dimension Function is almost constant in each range, there is a sharp contrast between the two ranges. This is characteristic of a bifractal. The observed step in the plot at about $k = 50$ represents the average number of luminous matter in a cluster. For $\gamma < -5$ as plotted in Fig. 5.12, the large step should not be some numerical artifacts as $k = 50$ is large enough to avoid the singularities from the gamma function in Eq. (3.7). the However, due to the smaller steps around $k = 50$, it is difficult to say if the large step at $k = 50$ is part of the structure within a fractal or the boundary of the scaling region.

N=10000 T=20 (Dark)

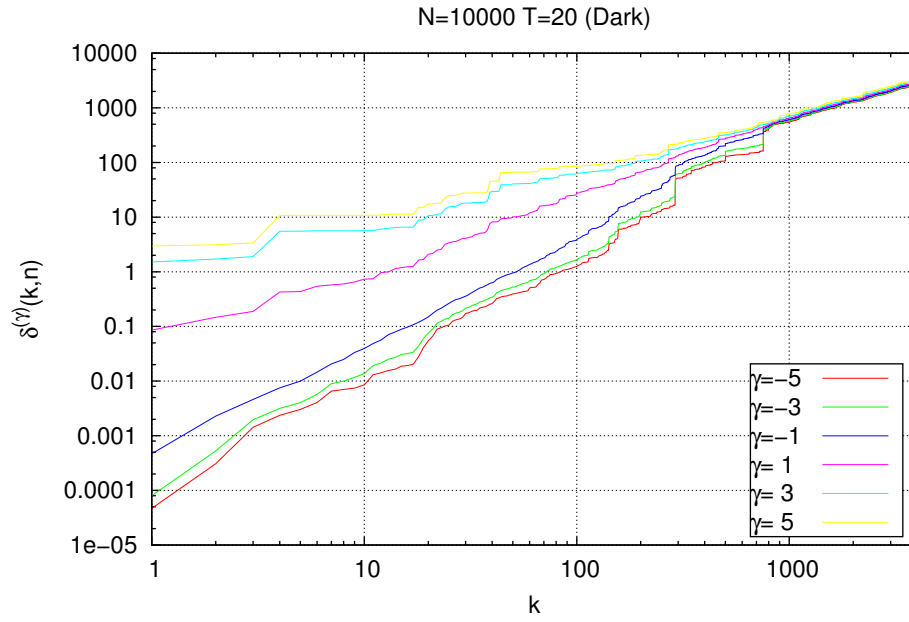

FIGURE 5.13: **The scaling of the kth neighbor distribution with dark matter at** $T = 20$.

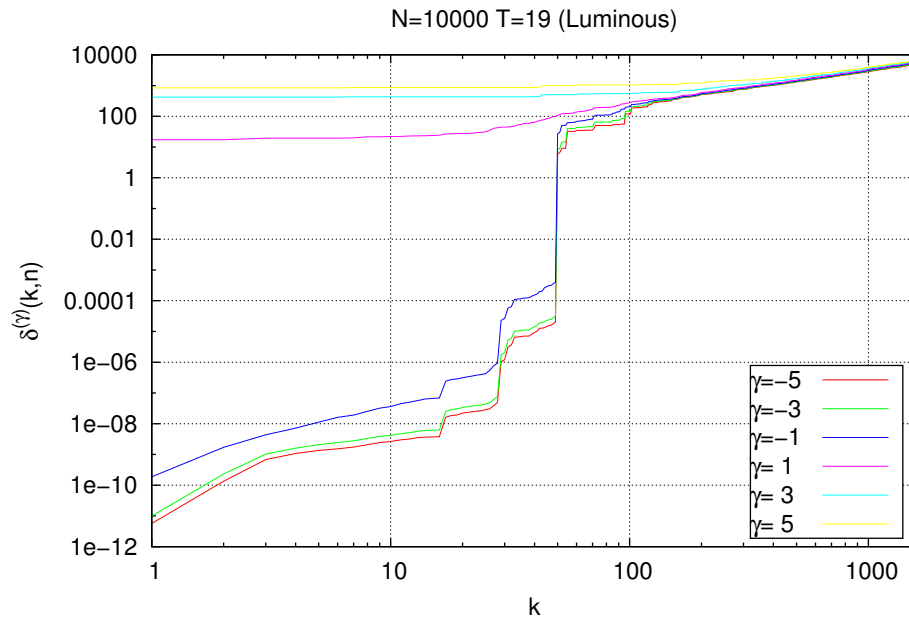N=10000 T=19 (Luminous)

FIGURE 5.14: **The scaling of the kth neighbor distribution with luminous matter at** $T = 19$.

5.7   Conclusion

While our one-dimensional model does not make quantitative predictions at this point, it gives qualitative insights regarding the evolution of structure formation in the actual universe. We have found that our one-dimensional model is qualitatively consistent with the standard scenario of structure formation: at early times, the mass distribution was taken to be homogenous across the space, following the theoretical distribution of the primordial universe. We observed that the matter-particles form clusters with large voids between them as the time increases, both in the configuration space and in the μ space. The probability distribution plots at a later time T show that, at the core of the clusters, we have a concentration of luminous matter. On the other hand, dark matter is more evenly spread out around the cluster cores. While bigger clusters contain both dark matter and luminous matter, beyond a certain resolution, we have also found smaller clusters which are entirely consisting of luminous matter or dark matter. This suggests that the algorithm is no longer a good indicator of what is actually happening in the universe beyond this resolution.

To investigate the fractal nature of the clusters, we have applied mass-oriented methods to the matter-particles in our model. Especially, the k-neighbor method, the global method employing a partition of equal mass, proves to be useful in investigating the evolution of the large-scale structure formation. Our results show that the beyond certain k at each T, the matter distribution scales homogeneously, with the associated dimension of 1. We found that this lower bound of homogeneity increases, meaning that the scale beyond which the universe is homogenous expands over time. This observation confirms the bottom-up structure formation and the cosmological principle simultaneously.

Below the point of homogeneity, we have a large enough scaling to extract the associated multifractal dimensions. This means that the simulated set of matter-particles is creating hierarchical patterns which are difficult to replicate in three-dimensional models. The generalized dimensions from these scaling ranges show that the simulated set is highly multifractal with the box-counting dimension $D_0 = 0.7$ well below the embedded dimension. The spectra of the generalized dimensions are distinctively different for two species of matter. As luminous matter tends to lose its energy faster once it is trapped in a cluster, luminous

mater subsequently creates more lacunarity in the set. Consequently, the mass distribution of luminous matter resembles that of a bifractal. On the other hand, the mass distribution of dark matter has a typical multifractal spectrum.

5.8   Further Research

There are several paths we want to pursue using this one-dimensional model. First, we would like to investigate how structure formation changes as the initial parameters are varied. The key parameters include the initial primordial fluctuation distribution, the rate of energy loss and the dark matter to luminous matter ratio. Some argue that the CMB cold spot may be explained by non-Gaussian primordial fluctuations [47]. Therefore, it would be interesting to see how the initial distribution affects the evolution of matter. The dark to luminous matter ratio is also a subject of debate. Since the dark matter is not directly observable, the estimate of the dark matter density may be biased.

Second, comparing the fractal dimension between different clusters, and at different degrees of hierarchy, should provide some insights into the bottom-up structure formation. How to define a cluster in a fractal geometry is also the subject of our future research. It should also be beneficial to compute the fractal dimension in phase space as well since two independent clusters crossing each other cannot be distinguished in configuration space alone.

Finally, we would like to investigate the reliability of this one-dimensional model. We can invert the transformation from the comoving frame back to the original space-time variables and see if the model is consistent with some of the theoretical predictions of present day cosmology.

Appendix A

Mathematics

This appendix contains minimal explanations for the mathematical concepts introduced in Chapter 2. For more detail, see any standard analysis textbook [48]. Here we mainly follow the formulations given in *Fractal Geometry: Mathematical Foundations and Applications* by Kenneth Falconer [14].

A.1 Measure

In mathematics, the concept of measures is introduced to assign a "size" to a given set. The idea of measures may be implicitly used when we speak of the length, area or volume of a set. For this reason, the definition of measures is constructed in a way that extends our conventional notion of size. Precisely, a measure μ on $\mathbb{R}^n$ is a function which assigns a non-negative number to subsets A and B of $\mathbb{R}^n$ such that the following properties hold:

  i $\mu(\varnothing) = 0$

 ii $\mu(A) \leq \mu(B) \quad$ if $A \subset B$

iii If $A_1, A_2, \ldots$ is a countable sequence of sets, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i) \tag{A.1}$$

The first condition guarantees that the measure of an empty space is zero. The second condition means that a measure is monotonic. It roughly says that the larger the set, the the larger the measure. The third condition can be simplified if the equality holds. The equality in Eq. (A.1) holds for a sequence of disjoint Borel sets $\{A_i\}$. It then states that a measure is countably additive.

A Borel set is formed through the operations of the countable union, countable intersection and taking complements of open sets or closed sets. It is important to notice that the definition of measures does not eliminate the existence of non-measurable sets. Borel sets are important since any measure defined on an open set, or closed set, must also be defined on all Borel sets.

A.1.1 Lebesgue Measure

The Lebesgue measure is the most commonly used measure. Unless otherwise explicitly stated, the Lebesgue measure is often assumed in literature when discussing the "size" of a given set. First, the Lebesgue outer measure is defined for a set $F \subset \mathbb{R}$ as:

$$\mu^* = \inf \left\{ \sum_{i=1}^{\infty} \ell(I_i) : \{I_i\} \text{ a sequence of intervals with } F \subseteq \bigcup_{i=1}^{\infty} I_i \right\} \tag{A.2}$$

where the length $\ell(I)$ of a bounded interval I whose endpoints a and b with $a < b$ is given by $\ell(I) = b - a$. Then the set F has the Lebesgue measure $\mu(F) = \mu^*(F)$ if $\forall$ set $A \subset \mathbb{R}$, it satisfies the following:

$$\mu^*(A) = \mu^*(A \cup F) + \mu^*(A \cap F^c). \tag{A.3}$$

For example, the set of rational numbers in a interval has Lebesgue measure zero as it is countable. The Cantor set also has Lebesgue measure zero.

Here, we follow the standard notation for infima. The infimum of a set S is the greatest lower bound of set S. Formally, the infimum $\inf(S)$ of a subset S of a partially ordered set P is a unique element $l \in P$ such that the following properties hold true:

i $l \leq x \ \forall x \in S$

ii $\forall y \in P$, if $y \leq x \ \forall x \in S$, then $y \leq l$.

Note that $l = \inf(S)$ may not be in the subset S. The completeness of real number guarantees the existence of such element for a subset bounded from below. The supremum of a set S or the least upper bound of a set S can be defined in a similar manner and is denoted by $\sup(S)$.

A.2 Cover

A cover plays a crucial role in fractal analysis. Oftentimes, it is the way a cover is constructed that leads to different definitions of fractal dimensions and associated methods for obtaining them. A countable collection C of sets $\{U_i\}$ is called a cover of set F if

$$F \subseteq \bigcup_{i=1}^{\infty} U_i. \tag{A.4}$$

In this thesis, we introduced a specific cover called δ-cover. A cover C is a δ-cover of F if it satisfies

$$0 < |U_i| \leq \delta \tag{A.5}$$

for each i. Note that here we adopted the notation $|U|$ for the diameter of a set U, which is given by

$$|U| = \sup\{|x - y| : x, y \in U\}. \tag{A.6}$$

We frequently say that C *covers* F when C is a cover of a set F.

Bibliography

[1] Clifford Pickover. *The math book : from Pythagoras to the 57th dimension, 250 milestones in the history of mathematics.* Sterling Pub, New York, NY, 2009. ISBN 1402757964.

[2] Benoit Mandelbrot. *The fractal geometry of nature.* W.H. Freeman, San Francisco, 1982. ISBN 0-7167-1186-9.

[3] L Pietronero. The fractal structure of the universe: correlations of galaxies and clusters and the average mass density. *Physica A: Statistical Mechanics and its Applications*, 144(2):257–284, 1987.

[4] Ji-sheng LIU and Yan-guang CHEN. Fractal studies of urban geography in the past and future [j]. *Scientia Geographica Sinica*, 2:012, 2000.

[5] P. Hasal and A.F. Munster. Spatiotemporal chaos in an electric current driven ionic reaction-diffusion system. *Chaos*, 4(3):531, 1994. ISSN 10541500. URL http://lib.tcu.edu/PURL/EZproxy_link.asp?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=cookie,ip,uid&db=a9h&AN=4200144&site=ehost-live&scope=site.

[6] Apostolos Serletis and Ioannis Andreadis. Random fractal structures in north american energy markets. *Energy Economics*, 26(3):389–399, 2004.

[7] Yurij Baryshev. *Discovery of cosmic fractals.* World Scientific, River Edge, N.J, 2002. ISBN 978-9810248710.

[8] R Brent Tully, Hélène Courtois, Yehuda Hoffman, and Daniel Pomarède. The laniakea supercluster of galaxies. *Nature*, 513(7516):71–73, 2014.

[9] Jean-Michel Alimi, Vincent Bouillot, Yann Rasera, Vincent Reverdy, Pier-Stefano Corasaniti, Irene Balmes, Stéphane Requena, Xavier Delaruelle, and Jean-Noel Richet. Deus full observable ΛCDM universe simulation: the numerical challenge. *arXiv preprint arXiv:1206.2838*, 2012.

[10] Luigi Guzzo, Angela Iovino, Guido Chincarini, Riccardo Giovanelli, and Martha P Haynes. Scale-invariant clustering in the large-scale distribution of galaxies. *The Astrophysical Journal*, 382:L5–L9, 1991.

[11] Vicent J Martínez and Bernard JT Jones. Why the universe is not a fractal. *Monthly Notices of the Royal Astronomical Society*, 242(4):517–521, 1990.

[12] Yui Shiozawa, Bruce N Miller, and Jean-Louis Rouet. Fractal dimension computation from equal mass partitions. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 24(3):033106, 2014.

[13] Helge Von Koch. On a continuous curve without tangent constructible from elementary geometry. *Classics on Fractals (Westview Press, 2004) pp*, 25:45, 1904.

[14] K. J. Falconer. *Fractal geometry mathematical foundations and applications*. Wiley, Chichester, 2003. ISBN 978-0470848623.

[15] Patricia M. E. Altham. Two generalizations of the binomial distribution. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 27(2):pp. 162–167, 1978. ISSN 00359254. URL http://www.jstor.org/stable/2346943.

[16] Alfréd Renyi. Dimension, entropy and information. In *Trans. 2nd Prague Conf. Information Theory*, pages 545–556, 1960.

[17] R. Riedi. An improved multifractal formalism and self-similar measures. *Journal of Mathematical Analysis and Applications*, 189(2):462 – 490, 1995. ISSN 0022-247X. doi: http://dx.doi.org/10.1006/jmaa.1995.1030. URL http://www.sciencedirect.com/science/article/pii/S0022247X8571030X.

[18] Jens Feder. *Fractals*. Plenum Press, New York, 1988. ISBN 0306428512.

[19] Remo Badii and Antonio Politi. Statistical description of chaotic attractors: The dimension function. *Journal of Statistical Physics*, 40(5-6):725–750, 1985. ISSN 0022-4715. doi: 10.1007/BF01009897. URL http://dx.doi.org/10.1007/BF01009897.

[20] Willem van de Water and Piet Schram. Generalized dimensions from near-neighbor information. *Phys. Rev. A*, 37:3118–3125, Apr 1988. doi: 10.1103/PhysRevA.37.3118. URL http://link.aps.org/doi/10.1103/PhysRevA.37.3118.

[21] G Broggi. Evaluation of dimensions and entropies of chaotic systems. *JOSA B*, 5(5): 1020–1028, 1988.

[22] Tamás Tél, Ágnes Fülöp, and Tamás Vicsek. Determination of fractal dimensions for geometrical multifractals. *Physica A: Statistical Mechanics and its Applications*, 159 (2):155–166, 1989.

[23] Renaud Lopes and Nacim Betrouni. Fractal and multifractal analysis: a review. *Medical image analysis*, 13(4):634–649, 2009.

[24] William Press. *Numerical recipes : the art of scientific computing*. Cambridge University Press, Cambridge, UK New York, 2007. ISBN 0521880688.

[25] Milton Abramowitz. *Handbook of mathematical functions : with formulas, graphs, and mathematical tables*. Dover Publications, New York, 1970. ISBN 978-0486612720.

[26] William Feller. *An introduction to probability theory and its applications*. Wiley, New York, 1968. ISBN 978-0471257080.

[27] R Badii and G Broggi. Measurement of the dimension spectrum $f(\alpha)$: Fixed-mass approach. *Physics Letters A*, 131(6):339–343, 1988.

[28] Peter Grassberger. Generalizations of the hausdorff dimension of fractal measures. *Physics Letters A*, 107(3):101–105, 1985.

[29] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173, 1929.

[30] Scott Burles, Kenneth M. Nollett, and Michael S. Turner. Big bang nucleosynthesis predictions for precision cosmology. *The Astrophysical Journal Letters*, 552(1):L1, 2001. URL http://stacks.iop.org/1538-4357/552/i=1/a=L1.

[31] JR Bond and G Efstathiou. The statistics of cosmic background radiation fluctuations. *Monthly Notices of the Royal Astronomical Society*, 226(3):655–687, 1987.

[32] Volker Springel, Carlos S Frenk, and Simon DM White. The large-scale structure of the universe. *Nature*, 440(7088):1137–1144, 2006.

[33] Volker Springel, Simon DM White, Adrian Jenkins, Carlos S Frenk, Naoki Yoshida, Liang Gao, Julio Navarro, Robert Thacker, Darren Croton, John Helly, et al. Simulations of the formation, evolution and clustering of galaxies and quasars. *nature*, 435 (7042):629–636, 2005.

[34] James Rich. *Fundamentals of cosmology*. Springer, Berlin, 2009. ISBN 978-3642027994.

[35] Phillip James Edwin Peebles. *Principles of physical cosmology*. Princeton University Press, 1993.

[36] L Pietronero. The fractal structure of the universe: correlations of galaxies and clusters and the average mass density. *Physica A: Statistical Mechanics and its Applications*, 144(2):257–284, 1987.

[37] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: Evidence, candidates and constraints. *Physics Reports*, 405(5):279–390, 2005.

[38] Douglas Clowe, Maruša Bradač, Anthony H Gonzalez, Maxim Markevitch, Scott W Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *The Astrophysical Journal Letters*, 648(2):L109, 2006.

[39] J Dunkley, E Komatsu, MR Nolta, DN Spergel, D Larson, G Hinshaw, L Page, CL Bennett, B Gold, N Jarosik, et al. Five-year wilkinson microwave anisotropy probe observations: likelihoods and parameters from the wmap data. *The Astrophysical Journal Supplement Series*, 180(2):306, 2009.

[40] Liang Gao and Simon DM White. Assembly bias in the clustering of dark matter haloes. *Monthly Notices of the Royal Astronomical Society: Letters*, 377(1):L5–L9, 2007.

[41] James M Bardeen, Paul J Steinhardt, and Michael S Turner. Spontaneous creation of almost scale-free density perturbations in an inflationary universe. *Physical Review D*, 28(4):679, 1983.

[42] Tomoaki Ishiyama, Keigo Nitadori, and Junichiro Makino. 4.45 pflops astrophysical n-body simulation on k computer–the gravitational trillion-body problem. In *High Performance Computing, Networking, Storage and Analysis (SC), 2012 International Conference for*, pages 1–10. IEEE, 2012.

[43] J.-L. Rouet, M. R. Feix, and M. Navet. Fractals in astronomy. In A. Heck, editor, *Vistas in Astronomy 33*, pages 357–370. Pergamon, 1990.

[44] Bruce N Miller and Jean-Louis Rouet. Cosmology in one dimension: fractal geometry, power spectra and correlation. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(12):P12028, 2010.

[45] Bruce N Miller and Jean-Louis Rouet. Ewald sums for one dimension. *Physical Review E*, 82(6):066203, 2010.

[46] James Binney. *Galactic dynamics*. Princeton University Press, Princeton, 2008. ISBN 9781400828722.

[47] Marcos Cruz, L Cayón, E Martinez-Gonzalez, P Vielva, and J Jin. The non-gaussian cold spot in the 3 year wilkinson microwave anisotropy probe data. *The Astrophysical Journal*, 655(1):11, 2007.

[48] Walter Rudin. *Principles of mathematical analysis*. McGraw-Hill, New York, 1976. ISBN 978-0070542358.

Yui Shiozawa

Personal Data

Place and Date of Birth:	Japan — 23 February 1986
Family:	Daughter of Yoshinori and Shizuko Shiozawa
Address:	2801 Princeton St., Fort Worth, TX, USA
Phone:	+1 903 386 5005
email:	yui.shiozawa@tcu.edu

Education

June 2010 Bachelor of Science with Honors
with *Philosophy and Physics Specialist* and
with *Mathematics Major*
The University of Toronto, Toronto, ON, Canada

June 2004 Secondary School Diploma
Bishop's College School, Sherebrook, QC, Canada

Experience

2012-2014 Teaching Assistantship
Texas Christian University, Fort Worth, TX, USA

Professional Memberships

American Physical Society

ABSTRACT

FRACTAL ANALYSIS
IN
A ONE-DIMENSIONAL UNIVERSE

by Yui Shiozawa, M.Sc., 2014
Department of Physics and Astronomy
Texas Christian University

Thesis Advisor: Bruce N. Miller, Professor of Physics
                  Hana M. Dobrovolny, Assistant Professor of Biophysics
                  Peter M. Frinchaboy, Assistant Professor of Astrophysics
                  Igor Prokhorenkov, Associate Professor of Mathematics
                  Yuri M. Strzhemechny, Associate Professor of Physics

While the universe we observe today exhibits local filament-like structures, with stellar clusters and large voids between them, the primordial universe is believed to have been nearly homogeneous with slight variations in matter density. To understand how the observed hierarchical structure was formed, researchers have developed a one-dimensional analogue of the universe that can simulate the evolution of a large number of matter particles. Investigations to date demonstrate that this model reveals structure formation that shares essential features with the three-dimensional observations. In the present work, we have expanded on this concept to include two species of matter, specifically dark matter and luminous matter. In our simulation, luminous matter is treated in a way that loses energy in interaction. The results of the simulations clearly show the formation of a Cantor set like multifractal pattern over time. In contrast with most earlier studies, mass-oriented methods for computing multifractal dimensions were applied to analyze the bottom-up structure formation.