PROBING THE NON-LINEARITY IN GALAXY CLUSTERS THROUGH THE ANALYSIS OF FRACTAL DIMENSION VIA WAVELET TRANSFORM

Loay Ali Khalifa

DePaul, Khalifa_loay@yahoo.com

Recommended Citation

https://via.library.depaul.edu/csh_etd/273
PROBING THE NON-LINEARITY IN GALAXY CLUSTERS THROUGH THE ANALYSIS OF FRACTAL DIMENSION VIA WAVELET TRANSFORM

A Thesis
Presented in
Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE

August, 2018

BY
Loay A. M. Khalifa

PHYSICS DEPARTMENT
College of Science and Health
DePaul University
Chicago, Illinois
# TABLE OF CONTENTS

**LIST OF FIGURES** .................................................. 5  
**LIST OF TABLES** ................................................... 8  
**ABSTRACT** ........................................................... 9  
  0.1 Introduction .................................................... 10  
  0.2 Acknowledgment ................................................ 13  
  0.3 Dedication ....................................................... 16  

**CHAPTER 1 Fractals in Cosmology** ............................... 17  
  1.1 The Universe Through History ................................ 17  
  1.2 The Universe According To Modern Cosmology ............... 21  
    1.2.1 The Cosmological Principle ............................... 21  
    1.2.2 The Universe on Large Scales ........................... 23  
    1.2.3 Dark Energy ............................................... 28  
  1.3 Statistics of Matter Clustering on Large Scales ............ 30  
    1.3.1 Density Field .............................................. 31  
    1.3.2 Two Point Correlation Function .......................... 31  
    1.3.3 Power Spectrum ............................................ 33  
  1.4 Fractals in Cosmology ......................................... 34  
    1.4.1 On The Meaning of Being A Fractal ...................... 34  
    1.4.2 Examples of Fractals ...................................... 36  
    1.4.3 Fractal Analysis of Galaxies ............................. 37  
  1.5 The Sloan Digital Sky Survey SDSS ............................ 39  
    1.5.1 Baryonic Oscillations Spectroscopic Survey BOSS ....... 40
TABLE OF CONTENTS – Continued

1.5.2 Mock Galaxy Catalogs ...................................... 41
1.6 Goals and Motivations ......................................... 42

CHAPTER 2 Wavelets ............................................... 44
2.1 Definition and Properties ....................................... 44
  2.1.1 Haar Wavelet .............................................. 45
2.2 Wavelet Transforms ............................................. 47
  2.2.1 The Discrete Wavelet Transform ......................... 48
  2.2.2 The Wavelet Packet Transform ......................... 53
2.3 2-D Wavelet Transforms ........................................ 55
2.4 Why Wavelets? .................................................. 57

CHAPTER 3 Estimating the Fractal Dimension $D$ .............. 60
3.1 Workflow .......................................................... 60
3.2 Reproducing the Distribution of Galaxies ..................... 61
  3.2.1 Raw SDSS Data ............................................ 62
  3.2.2 Creating the Matrices using HealPix and SDSSPix ....... 62
3.3 Calculating the Power Spectrum ................................ 67
  3.3.1 Estimating the Power Spectrum From Wavelets ......... 67
  3.3.2 Frequency of the Signal .................................... 69
  3.3.3 Radial Averaging .......................................... 74
3.4 Estimating the Hurst Exponent $H$ ............................ 76
  3.4.1 On the Meaning of the Hurst Exponent ................. 76
  3.4.2 Estimating the Value of the Hurst Exponent .......... 77
3.5 Calculating The Fractal Dimension ............................ 80

CHAPTER 4 Results and Interpretations .......................... 82
4.1 Results ............................................................ 82
TABLE OF CONTENTS – Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2 Interpretation</td>
<td>94</td>
</tr>
<tr>
<td>CHAPTER 5 Conclusions and Future Research Goals</td>
<td>96</td>
</tr>
<tr>
<td>5.1 Conclusion</td>
<td>96</td>
</tr>
<tr>
<td>5.2 Goals of Future Research</td>
<td>97</td>
</tr>
<tr>
<td>Appendices</td>
<td>99</td>
</tr>
<tr>
<td>APPENDIX A Program: Powerspectrum</td>
<td>100</td>
</tr>
<tr>
<td>APPENDIX B Program: Datahandler</td>
<td>104</td>
</tr>
<tr>
<td>APPENDIX C Program: myhealpix</td>
<td>105</td>
</tr>
<tr>
<td>APPENDIX D Program: mymatrix</td>
<td>107</td>
</tr>
<tr>
<td>APPENDIX E Program: PropDist</td>
<td>109</td>
</tr>
<tr>
<td>APPENDIX F Program: Results</td>
<td>110</td>
</tr>
</tbody>
</table>


LIST OF FIGURES

1.1 Al-Sufi’s depiction of Orion as seen from the Earth from his book: The Book of Fixed Starts. The two illustrations are mirror images of each other. Every red dot represents an object in the sky, and the long sleeve is a representation of Orion’s shield. This image dates back to the 13th century. Source: reproduced from Bibliothèque National de France. ................................................................. 20

1.2 Thompson and Gregory’s figure from their 1978 paper. The figure shows the relation between the right ascension, distance in Mpc, and redshift. Each one of the black dots represents a galaxy, and these galaxies together are grouped into clusters, which are grouped into superclusters. The existence of significant voids is clear, and this structure repeats itself. The Milky Way galaxy is located at the bottom of this slice [32]. ................................................................. 23

1.3 Depiction of the main events occurred during the entire history of the Universe. Source: reproduced from the European Southern Observatory ESA. Image Courtesy: ESA C. Carreau. ................................. 28

1.4 Illustration of the Von Koch Curve F. The first panel on the top is just a straight line labeled by $E_0$ as the first stage in constructing the Koch Curve. At the next panel $E_1$, the straight line is partitioned into three segments, and the middle third section is replaced by the other two sides of an equilateral triangle. At subsequent stages of the construction, the same process of partitioning is repeated until the Koch Curve $F$ is produced at the lower most panel [11]. .......................... 37

1.5 The SDSS coverage. The BOSS survey covers the redshift range $0 < z < 0.8$. Source: reproduced from the SDSS website. ......................... 41

2.1 The Haar Scaling function $\phi(r)$ along the interval (0,1). The function is equal to zero outside the interval [0,1). Source: reproduced from Texas A and M University Math 414 - Spring 2004 Course Website. .............................. 46

2.2 The Haar wavelet $\psi(r)$ along the interval (0,1). Source: reproduced from Wikimedia Commons, the free media repository. ..................... 47
LIST OF FIGURES – Continued

2.3 The act of the low and high pass filters on a simulated signal of a sine wave with noise. .......................................................... 49

2.4 The DWT tree. The original signal in the top is transformed into the approximation and detail coefficients after the first pass or sweep \( j = 1 \). Then at level \( j = 2 \), we perform the DWT again only on the approximation coefficients, and the detail coefficients are left without any more transformation. We proceed further with more passes or levels following the same procedures. ........................................ 52

2.5 The WPT tree. The vector \( S \) is the original signal. After the first pass or sweep \( j = 1 \), the original signal vector is transformed into approximation coefficients \( a_i \) and detail coefficients \( d_i \). The approximation coefficients are the result of the action of the low pass filter on the signal. The detail coefficients are the result of the action of the high pass filter on the signal. ...................................................... 55

2.6 The two dimensional Wavelet Packet Transform Tree. The low and high pass filters acts on all types of coefficients at each pass. Each set of coefficients is represented by a matrix [37]. ........................................ 57

2.7 Comparison between the WPT and DWT. In the case of DWT, one branch of the leaves is terminated after the first pass so the leaves have only one branch. For the WPT, the leaves do not terminate at any resolution. As a result, we end up with more and narrower frequency bands to examine the signal. The sweep number is \( j \), and \( a \) and \( d \) are the approximation and detail coefficients, respectively. ... 59

3.1 Workflow of the entire process. .......................................................... 61

3.2 The celestial sphere is divided into pixels of equal sizes. Depending on the desired resolution, the partitioning of the sphere is varied into low resolution or high resolution grids. The sphere on the upper left has the lowest possible resolution with 12 pixels of equal sizes and equal areas. As we move clockwise, the resolution is higher and the number of pixels is larger. The resolution of the sphere on the lower left is the highest with 768 pixels. Source: reproduced from the NASA Jet Propulsion Laboratory. https://healpix.jpl.nasa.gov ........................................ 63
LIST OF FIGURES – Continued

3.3 The two natural pixel indexing schemes, ring and nested. From top to bottom: Panel one (resolution = 2) and panel two (resolution = 4) show the ring scheme. Panel three (resolution = 2) and panel four (resolution = 4) show the nested scheme. Source: reproduced from the NASA Jet Propulsion Laboratory. https://healpix.jpl.nasa.gov

3.4 The natural ordering of the WPT tree leaves. The $H$ is the low pass filter and $G$ is the high pass filter. The down arrow followed by number 2 means that the number of coefficients is reduced by 2 after each pass. The $\chi_{i,j}$ notation represents the transformed coefficients obtained from the pass $i$, with a wavelet packet node number $j$ [21].

3.5 The WPT frequency tree for level of decomposition $n = 4$. Here $H(z)$ and $G(z)$ denote the low and high pass filters, respectively. The down arrows means that the number of coefficients is reduced by 2. The coefficients of the wavelet packet transform are ordered by the increasing order of frequency using the Gray Code permutation [21].

3.6 On the left, the plot of the power spectrum vs. the frequency obeys the power law distribution. To the right, the log-log plot of the power spectrum vs. the frequency, with line fitted to the points.

3.7 The power spectrum plotted against the physical scale obeys the power law distribution.

4.1 Hurst exponent at resolution 8.

4.2 Fractal dimension at resolution 8.

4.3 Hurst exponent at resolution 12.

4.4 Fractal dimension at resolution 12.

4.5 Hurst exponent at resolution 16.

4.6 Fractal dimension at resolution 16.

4.7 Hurst exponent at resolution 32.

4.8 Fractal dimension at resolution 32.

4.9 Hurst exponent at resolution 64.

4.10 Fractal dimension at resolution 64.
**LIST OF TABLES**

1.1 Results of the fractal Analysis of different surveys from the 1997 Pietronero, Montuori and Labini paper [7]. $D$ is the fractal dimension. 39

4.1 Slopes of BOSS and Mock Catalogs Galaxies at resolution 8 . . . . . 84
4.2 Slopes of BOSS and Mock Catalogs Galaxies at resolution 12 . . . . . 86
4.3 Slopes of BOSS and Mock Catalogs Galaxies at resolution 16 . . . . . 88
4.4 Slopes of BOSS and Mock Catalogs Galaxies at resolution 32 . . . . . 90
4.5 Slopes of BOSS and Mock Catalogs Galaxies at resolution 64 . . . . . 92
The study of large scale structure (LSS) of the Universe armed with both all-sky surveys and numerical simulations has become an increasingly important tool to probe basic cosmology. We used the method of wavelet transforms combined with the fractal based point-processes to investigate the clustering of matter on galactic scales through the fractal analysis approach. In particular, we developed a method to calculate the angular fractal dimension of galaxy distributions as a function of the cosmological comoving space. Taking advantage of the self-similarity and localization properties of the wavelets, allows us to compute the fractal dimension of galaxies in narrow redshift bins. The narrow bins assure that dynamical evolution has not occurred. We used both real and simulated data from the Baryon Oscillation Spectroscopic Survey (BOSS) and the Mock Galaxy Catalogs produced by the Sloan Digital Sky Survey (SDSS). Using the wavelet packet power spectrum, we find areas in the galaxy distribution which have power law like behavior. The exponent of the power law is the Hurst exponent $H$, which is directly related to the fractal dimension of spatial point processes. We find the fractal dimension at all redshifts is $D = 1.3 \pm 0.2$ for BOSS Galaxies while $D = 1.6 \pm 0.3$ for Mock Galaxy Catalogs. We concluded that galaxies distribution in the redshift range $z < 1$ can be described as angular fractal systems, and the distribution is inhomogenous and irregular.
0.1 Introduction

The rise of cosmology as a modern science in the 20th century is one of great scientific triumphs of humanity. The Universe, as we write these lines to the best of our understanding, is accelerating its expansion, is isotropic and homogenous in large scales over 150 Mpc, described by the Λ CDM model, and dominated by the dark energy. However, the Universe is inhomogenous on much smaller scales. The departure from inhomogeneity to homogeneity is significantly important in understanding the structure formation in the Universe and provides us with valuable tools to establish cosmological parameters. For example, the baryonic acoustic oscillations serve both to confirm the Λ CDM model and to fix basic parameters like the Hubble constant. Furthermore, the growth of structure as seen in the galaxy evolution provides us with tools to dismantle the expansion history of the Universe, so we can understand more about the nature of the mysterious dark energy.

In this thesis, we develop a new method to calculate the angular fractal dimension of galaxy clusters using wavelet transform methods on the near Universe ($z < 1$). In particular, we calculate the fractal dimension as a function of cosmological comoving space over narrow-bins for galaxies. We use data from the SDSS BOSS survey and SDSS Mock Galaxy Catalogs. Chapter 1 in the thesis serves as the background of this project. We introduce the basic cosmology we used in our work - the theory of large scale structure. We discuss what fractals are and the previous work done on the fractal analysis of galaxies. We also explore the nature of the data we used - the Baryonic Oscillations Spectroscopic Survey (BOSS) and the Mock Galaxy Catalogs of the Sloan Digital Sky Survey. We conclude the chapter with the goals and motivations of our research. In chapter 2, the main tools of the research will be introduced - mainly the wavelet transform methods. We will provide a detailed quantitative and qualitative discussion on wavelets and how we used the wavelet transform in our research. In chapter 3, we present the algorithm we developed to calculate the fractal dimension of galaxies and the various procedures we followed. In chapter 4, we will present our analysis of the data, the results, and our interpretation.
Finally, in chapter 5, we will present our conclusions and discuss potential future research.
“It is good for the soul to suffer, the more it suffers the stronger it is - the better in shape it is.”

Jesús Pando
0.2 Acknowledgment

All praise is due to Almighty ALLAH who bestowed His grace upon me to complete this thesis and the M.Sc degree requirements, it would have been impossible to reach this without His Blessings. It is hard to believe that my Fulbright journey at DePaul come to an end which makes me feels both happy and sad. On one hand, happy to get the M.Sc in Physics without a B.Sc in Physics. On the other hand, I am sad because I am leaving the Physics Department which was my home in the last two years. I will not forget the wonderful time I spent here and the people I had the pleasure to meet at Byrne Hall throughout my life.

Back in June 2013 after I earned my B.Sc degree in Geophysics from Ain Shams University in Egypt, I would not have imagined, even in my wildest dreams, that I will at some day in the future, earn an M.Sc in Physics form DePaul University and write and defend a thesis about cosmology. A long list of people have made this possible.

First and foremost, Dr. Jesús Pando, my thesis adviser, deserves the biggest gratitude for having the patience to have me as his student in the last two years. I had 4 classes and 2 thesis research quarters with him. The PHY 442 class was my first interaction with programming and Dr Pando has had the patience to explain to me what seemed to others rudimentary. In Winter Quarter 2018, I had dominated Dr. Pando’s schedule with PHY 412 and PHY 475. Perhaps the PHY 475 Cosmology class was the best class I have ever had; indeed, this class was the real joy. When it comes to the thesis, I knew very little about computation and nothing about cosmology when I started this research. I learned most of what I know now from our long, serious, funny, and instructive discussions. It would sounds more fair to say this thesis was prepared by Jesús Pando and Loay Khalifa - for the contribution of Jesús in this research. Whenever the word “we” is mentioned in this thesis, it must be interpreted as “Loay Khalifa under the unconditional guidance of Jesús Pando.”

I had a meeting with him virtually everyday for the last year either to fix my codes, to ask about cosmology, and seeking advice from him. Dr. Diva was always gen-
erous in his time despite being always busy as the Department Chair, an adviser, a teacher, a university leader, and a frequent attendee of meetings for important people. He coined the famous quote, “I am doing you a favor by letting your soul suffer.” I could not having imagined to have a better adviser for my thesis. By all means, Mucho Gracias Senor Jesús - you made a big difference in my life.

I want to express my gratitude for all my family members for supporting me financially and emotionally during the last couple of years. A big “shukran” is for my mother, Eman, my father, Ali, my brother, Mohammad, and my sister, Mayar.

I am grateful for Dr. Sarma and Dr. Beck-Winchatz to serve in my thesis committee. Dr. Sarma coined the famous quote “this is going to be the most difficult class you will ever have in your life,” in referring to Jackson’s Electricity and Magnetism classes. The highs and lows of those classes made me a better student. Dr. Sarma was always helpful during the GRE Physics and grad school applications, so I would like to express my gratitude to him. Dr. Beck-Winchatz deserves a very big “Danke” for his help during the GRE Physics season, suggestions about grad school applications, and his joyful spirit in PHY 440.

I must express my gratitude to the US Department of State for funding my studies throughout the Fulbright Fellowship. It was the generous financial support of the Fulbright Foreign Student Program that allowed me to get an admission to DePaul. I am grateful for the Binational Fulbright Commission in Egypt Staff and the Fulbright Staff at AMIDEAST Washington, DC. Dianne Price, Brian Dilehay, Shanoon Conheady, Reem El-Hasan, and Iman Jamal worked tirelessly with me throughout the last three years. Thank you also for the folks at Education USA Egypt who directed me along the path when I started the process of pursuing a graduate degree back in 2013.

I am deeply indebted to all the people in the Physics department. I am grateful for Dr. Kustusch for being my instructor for three classes in the first year and particularly for the PHY 310 class which was my first Physics class after 5 years. I had dominated her office hours in the academic year 2016 - 2017 and she was always generous in her time, supportive, and helpful. She coined the term “You have a
different coordinate system,” referring to my struggle with PHY 310.
I would like to express my gratitude for Dr. Landahl our graduate program director who coined the other term “the long suffering international student,” referring to my struggle with PHY 470. I had known Dr Landahl since I wrote to him my first email back in August 2015, and he was always helpful and supportive whenever needed. The trio Dr. Sarma, Dr. Landahl, and Dr. Gonzalez, deserves a big thanks for interviewing me back in May 2016 and admitting me to the program. I have not had any feeling of joy similar to that I had when I received the admission email from Dr Landahl on Tuesday, May 10th, 2016.
Mary Ann the Department Assistant deserves a great gratitude for playing the role of my mother in the last two years. Thank you for all the food, the desert, the treats, and for taking care of me. Your generosity is rare.
My fellow undergraduate and later graduate students, Mike Shah and Justin Potvin, it has been a pleasure to meet both of you. Both of you played a pivotal role in my survival in first year of the program, so I am grateful for all of what you have done with me.
Thank you to my extended family and friends who supported me unconditionally throughout the last years.
Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions.
0.3 Dedication

For my mother Eman and father Ali, who supports me unconditionally throughout my life, and without whose support writing this thesis would not have been possible, this is a small way of saying thanks.
CHAPTER 1

Fractals in Cosmology

“And Solomon inherited David. He said, O people, we have been taught the language of birds, and we have been given from all things. Indeed, this is evident bounty.”
Quran 27:16

In this chapter we give a description about the background cosmology. Cosmology is the study of the Universe as a whole as a physical system. We start the chapter by a brief history of the development of our view of the Universe through thousands of years. After that in section 1.3, we introduce the basic cosmology that is important to our research. Then, we introduce the theory of large scale structure (LSS) formation and how it can be used to constrain cosmological parameters. The main constituents of LSS are individual galaxies which are grouped into clusters of galaxies, which are then grouped into superclusters of galaxies. In section 1.4, we will introduce the reader to fractals and the machinery to calculate them. After that, we will explain what we are trying to measure which is the fractal dimension of galaxies, and present some previous results on the fractal analysis of galaxies. In section 1.5, we introduce the data we used in this project, the galaxy catalogs of the Sloan Digital Sky Survey (SDSS). We conclude the chapter by the goals and motivations of our research.

1.1 The Universe Through History

Astronomy is among the oldest of all natural sciences. Humans, throughout history from antiquity to modern day, have studied the Universe for different reasons;
whether it is religious, mythological, economical, and most importantly, scientific. Astronomy was developed by all civilizations in the east and west. Civilizations such as Mesopotamian, Chinese, Indian, Islamic, Ancient Greeks, Medieval Europe, Mayan, and Western civilizations had their equal share of contributions to the development of our view about the Universe.

The Ancient Greeks around the 4th century BC, believed that Earth is at the center of the Universe. They developed their model with Earth circled by the, Moon, and planets. Aristotle had argued that Earth is at the center of the Universe, and, Sun, Moon, planets, and stars are rotating around Earth. His theory was widely accepted at his time. Ptolemy proposed a geocentric model which stood as the standard theory of heavens for hundreds of years.

Indians, in the Indian subcontinent, studied astronomy during the Indus Valley Civilization around 3000 B.C.E. to create calendars to observe rituals. The oldest Indian astronomical text is the Vedanga Jyotisha which dates back to the Vedic period. The Aryabhatiya of Aryabhata is probably one of the most important texts in the history of Indian mathematics and astronomy. In China, astronomers developed the lunisolar calendar for purposes of time keeping, and they successfully made accurate predictions of eclipses. The polymath Su Song made significant contributions to the knowledge of stars, and his maps are the oldest survived manuscripts that showed the south polar projection. One of the earliest star catalogs was developed by Gan De in the 4th century BC.

In the Americas, the Mayans made accurate calculations of the Moon phases, the periodicity of eclipses, and the movement of Venus. Their most important and well-known contribution is the famous Mayan Calendar based on what then the Moon like other calendars of that time. They mastered the calculations of the cycles of the Pleiades, Sun, Moon, Venus, Jupiter, Saturn, and Mars to create their calendars which were vital to their rituals of the Mayan Religion.

Around 6 AD, during the rise of the golden age of Islamic Civilization, Muslim scientists translated the works of Greeks, Indians, and Persians to Arabic, and they
focused on improving the works of other civilizations. Under Caliph al-Mamun al-Rashid, the first observatory was built in Baghdad and another observatories were built in Iran. The telescope was not developed at that time, so astronomers between the 9th-11th AD centuries invented observational sextants. They used these tools to advance their knowledge about the Sun, the movement of stars, and orbits of planets. Abd al-Rahman al-Sufi known as (Azophi Arabus) published *Suwar Al-kawakib Al-thabita, The Book of Fixed Stars*. At that time, his text was the most comprehensive text on constellations in the sky. Al-Sufi made the first observation of the Andromeda galaxy and the Large Magellanic Cloud by naked eye. He marked it as a “cloud” in his notes because, of course, he did not realize that it was a galaxy. A figure from his work illustrates the Orion nebula is figure 1.1. A sufficiently large number of stars in the night sky, such as Aldebaran, Altair, and Deneb are still called today by their Arabic names. Some astronomical terms such as alidade, azimuth, and nadir, had their roots in Arabic.

Following the fall of the east civilizations, the Reconnaissance Age started in Europe. Johannes Kepler formulated his famous three planetary laws. Tycho Brahe made accurate and comprehensive planetary observations. Nicolaus Copernicus made, for the first time, accurate mathematical description of the heavens where he argued that the Sun is the center of the Universe. Galileo Galilei championed the Heliocentrism view of the Universe. Of course, the theories of the ancient civilizations were seriously flawed by the standards of modern science, but with the emergence of scientific revolution, astronomy became a modern science. Earlier in the 20th century, the seminal works of many scientists revolutionized our view of the Universe drastically. Harlow Shapley and Heber D. Curtis debated about the size of the Universe. Albert Einstein’s most famous work, General Theory of Relativity, provided a rigorous mathematical description of the Universe. Vesto Slipher made measurements of radial velocities of galaxies which suggested an expanding Universe. The evidence for an expanding Universe was independently suggested by Edwin Hubble and Georges Lemaitre. Lemaitre made the most crucial contributions
to the early stages of the Big Bang theory as an origin of the Universe. Lemaitre, Alexander Friedmann, Howard P. Robertson and Arthur Geoffrey Walker developed collaboratively the standard model of modern cosmology. The discovery of the Cosmic Microwave Background (CMB) in 1964 by Arno Penzias and Robert Wilson established the conclusive evidence of the big bang theory. By the end of the 90s, one of the most surprising discoveries in the history of science was found: that the expansion of the Universe is accelerating. The expansion is widely believed to be caused by the mysterious dark energy, or the famous Cosmological Constant. With the advent of powerful computers, new instruments, new ground and space based telescopes, and sophisticated technology, cosmology has entered the precision era. The development of cosmology as a modern science is one of the greatest scientific triumphs of humanity in the 20th century.

Figure 1.1: Al-Sufi’s depiction of Orion as seen from the Earth from his book: The Book of Fixed Starts. The two illustrations are mirror images of each other. Every red dot represent an object in the sky, and the long sleeve is a representation of Orion’s shield. This image dates back to the 13th century. Source: reproduced from Bibliotheque National de France.
1.2 The Universe According To Modern Cosmology

1.2.1 The Cosmological Principle

The local Universe is composed of stars, these stars combined together in the Milky Way galaxy. Next in the local neighborhood we see the Large Magellanic Cloud at 50 Kpc and the Andromeda, or M31 galaxy at a distance of 770 Kpc. Our galaxy resides within a small concentrated group of galaxies known as the Local Group of Galaxies, which composed of 30-40 galaxies with a diameter of around 3.1 Mpc. The two largest members of this group are the Milky Way and M31 (Andromeda). Further observations beyond 10 million light years we see around 20 groups of galaxies similar to our local group. Fifty million light years away from the Sun we find the nearest cluster of galaxies: the Virgo cluster. The Virgo cluster is composed of nearly 2000 small and 250 larger galaxies. Eventually, the clusters are grouped together into superclusters of galaxies. An example of that is the Local Supercluster of galaxies, where our Local Group and the Virgo cluster lies. This supercluster has a diameter of approximately 60 million light years. These layers of structures are observed everywhere in the Universe, so for cosmologists one crucial question is, what is the largest structure in the Universe? Or in other words, are superclusters the largest structures in our Universe?

To answer this question and other questions, galaxy surveys were conducted to build catalogs that were of essential importance to cosmologists in the later half of the 20th century. The primary goal of these surveys is mapping the spatial distributions of galaxies at different redshifts. The earliest work on this matter was the 50s efforts of Abell, Zwicky, and Lick. They successfully mapped the angular positions of around 1 million galaxies. Their catalogs, although marking a huge leap towards galaxy surveys, were not very helpful because they had no distance to galaxies information. A galaxy survey that would have the distance to galaxies was not achievable by the technological means of that time. However, during the late 50s, the 60s, and the 70s, the fundamental goal of observational cosmology was to design a galaxy
survey that would have redshift information. Abell, Zwicky, Schechtman, Gregory, Thompson, and the Harvard CfA led the efforts to design the redshift surveys. By the late 70s, the Harvard CfA finally designed the first red-shift survey. As more detailed and larger surveys have been designed, it became finally possible to answer the question: what is the largest structure in the Universe. Tifft and Gregory (1976) and Gregory and Thompson (1978) pioneering works on the large scale structure revealed the existence of significantly large voids between superclusters, as shown in figure 1.2. As seen from the figure, galaxies are grouped into clusters of galaxies, which are then grouped into superclusters. These supersclusters are separated by extremely large voids, and the combination of superclusters with voids is observed everywhere in the Universe. The observation of voids between superclusters was confirmed by a number of subsequent and more detailed redshift surveys like the 2dF redshift survey, and more recently, the SDSS survey.

Based on the observational evidence, it is obvious that these superclusters are separated by significantly large voids, and these are the largest scale structures in the Universe. These structures have been observed at scales at the level of 50 Mpc. On much larger scales of 100-150 Mpc, therefore, cosmologists assume that the Universe is, statistically speaking, isotropic and homogenous. That is, it looks the same everywhere (homogenous), and it looks the same in all directions (isotropic). Combined together, the homogeneity and isotropy assumptions are known as the Cosmological Principle (CP), and it is at the heart of modern cosmology. The importance of the Cosmological Principle to cosmology is similar to the importance of Newton’s Second law to classical mechanics, and to Schrodinger’s Equation to quantum mechanics. Modern cosmological observations such as the smoothness of the Cosmic Microwave Background radiation and the distribution of galaxies on large scales support the homogeneity and isotropy assumptions.
Figure 1.2: Thompson and Gregory’s figure from their 1978 paper. The figure shows the relation between the right ascension, distance in Mpc, and redshift. Each one of the black dots represents a galaxy, and these galaxies together are grouped into clusters, which are grouped into superclusters. The existence of significant voids is clear, and this structure repeats itself. The Milky Way galaxy is located at the bottom of this slice [32].

1.2.2 The Universe on Large Scales

On cosmological scales, the Universe is governed by the gravitational force, and dark energy plays a huge role, even greater than gravity, as we will see later. Gravitational instability is the basic driving force behind the formation of structures at large scales - clusters, superclusters, and voids [2]. In the following discussion, we provide a detailed, although a non-mathematical description of the main events that occurred during the last 14 billion years of the universe, and led to the current structure we observe.
From Inflation to the Cosmic Microwave Background

The events from inflation to the release of the Cosmic Microwave Background occurred between \( (t < 1 \text{ sec to } t = 380,000 \text{ years}) \). According to the inflation theory, the Universe underwent an extremely rapid exponential expansion which increased its volume when it was \( 10^{-32} \text{ seconds old} \). After the inflationary epoch, which ironed out most over densities and made space-time essentially flat, the Universe consisted of a uniform soup of fundamental particles including quarks, electrons, and their corresponding anti-particles. There were also neutrinos, photons or (particles of light), and dark matter particles. As a result of the continuous collision of the particles with their corresponding anti-particles, they annihilated each other. However, due to as yet understood reasons, there was slightly more matter than anti-matter. Therefore, the matter particles dominated the Universe, and the anti-matter is not present anymore.

At the same stage, quarks combined in trios to form protons. All these events happened in the immediate second after the Big Bang. When the Universe was around three minutes old after its creation, protons and neutrons had teamed up to form the hydrogen and helium nuclei. During this early time, the Universe had extremely high density and temperature. This period was characterized by the frequent collisions between the particles. In other words, the baryonic (ordinary) matter, electrons, protons, neutrons, and the few atomic nuclei that were present at that time were tightly coupled to the photons, using the language of cosmologists. As a result of the frequent interactions between baryons and photons, the latter were not capable of traveling freely. Hence, the Universe was opaque. Also, during that time, matter was under the influence of the gravitational force. Consequently, by the end of inflation, small density fluctuations that would be the seeds of structure later were present.

These over densities accumulated more mass as they attracted the matter from the neighborhood regions. At the same time, the baryons (ordinary matter) were still
coupled to the photons, and the radiation pressure of photons washed away the large accumulation of matter. It was radiation pressure that prevented the Universe from a total collapse under the effect of gravity. This continuous battle between the gravity and radiation pressure prevented the fluctuations to grow denser. The contest between the radiation pressure and gravity was responsible also to the propagation of the pressure or sound waves that would be later the Baryonic Acoustic Oscillations (BAO). During the same epoch, unlike the baryonic matter, the non-baryonic dark matter particles were not bound to the photons. Cosmologists refer to this type of dark matter particle as cold dark matter. This cold dark matter has a velocity that is much lower than the speed of light. Therefore, the density fluctuations of cold dark matter grew denser, and they became more massive. As time passed, these primordial fluctuations in the cold dark matter formed the seeds of the cosmic structure. The first objects that had formed in history had low mass, but they evolve eventually into more and more massive structures.

**Between the Cosmic Microwave Background to the Formation of the Earliest Stars and Galaxies**

During the Recombination epoch that lasted between \((t = 380,000 \text{ years to } t = \text{ a few hundred million years})\), the density of the Universe had lowered significantly in comparison to earlier times as a result of the continuous expansion. This expansion also caused the temperature to drop down to around 3000 Kelvin after it had reached billions of Kelvin in the immediate three minutes after the Big Bang. For the first time, atoms of neutral hydrogen formed after protons and electrons were finally able to combine. Besides that, decoupling of electrons and photons occurred, and cosmologists refer to this period by the name *Dark Ages*. The name reflects the fact that the Universe had no sources of light - it had only clouds made of neutral hydrogen. The results of decoupling were crucially important to the evolution of the Universe in the subsequent times. The first direct impact of decoupling was that for the first time, photons were capable of traveling across the Universe because they
were free from electrons. As they propagated throughout the entire Universe, the Universe became transparent. We observe this as the Cosmic Microwave Background (CMB). The second important effect of decoupling was that the baryonic matter began to aggregate under the influence of the gravity. This moment marked the beginning of the interaction of gravity with both the baryonic matter and non-baryonic dark matter. To that end, regions with higher concentrations of matter relative to their neighborhoods grew denser, and they became more massive.

Dark matter gravitational wells formed as a result of the inhomogeneities in the dark matter density field. The baryonic matter felt the attraction to these dense gravitational wells and eventually, these baryonic particles collapsed into the gravitational wells. Ordinary matter particles lost their energy by emitting radiation so they fell further and deeper into the dark matter potential wells. These dense regions became denser and more massive as this process continued, and eventually formed filaments of ordinary and dark matter and gave rise to the formation of the cosmic web. This complicated structure is the backbone of the development of stars and galaxies. Later, the regions of the highest matter concentration within the cosmic web gave rise to the first stars for the first time in the history of the Universe. These stars would later form galaxies, and the formation of the earliest stars eventually gave an end to the Dark Ages of the Universe.

After the Formation of the First Stars and Galaxies to Modern Day

During the time ($t = a$ few hundred million to $t = now$), the first stars appeared when the age of the Universe was a few hundred million years. The dense regions described above made the formation of stars and galaxies possible. The Universe consisted of stars and galaxies, and light from these structures traveled freely. This light is what the telescopes observe today, and it reveals the distant Universe to us as it was billions of years ago. These first stars were primarily composed of hydrogen and helium and they were 100 times or more massive than the Sun. However, their life was very short and they exploded into supernovae. These dying stars released
their remnants in the stellar neighborhoods which later gave birth to newer stars. Unlike the first stars that were made out primarily of hydrogen and helium, the composition of later generations of stars had other elements.

These elements were formed in the nuclear furnace of previous stars. The masses of the newer stars were typically smaller than their older counterparts. Surprisingly though, the first generation of galaxies were low-mass while the more massive galaxies formed later. The more massive structures like galaxy clusters and superclusters were formed later. This order of structure formation is what cosmologists observe today, and it is supported by the cold dark matter model. Figure 1.3 illustrates the main events that occurred during the age of the Universe: the last 14 billion years. These events starts in an increasing order of time from the left ($t=0$) to the right ($t=$present time). The events start with the Big Bang at time $t = 0$ and subsequent events like inflation, dark ages, formation of solar system occurred at different times. Of particular interest to our research is the time from Recombination about 380,000 years to the modern day. During this time, the structures in the Universe emerged from the earliest stars and galaxies to the superclusters of galaxies.

The quantitative description of this mechanism involves fairly complicated physics and math - enough to produce tens of dissertations. The interested reader is referred to the corresponding chapters at Ryden [1] and Dodelson [2] for a more in depth mathematical treatment of large scale structure formation.
1.2.3 Dark Energy

In the early 90s, the belief was that the Universe is full of matter and the attractive force pulls this matter together. As a result, the Universe had to slow down its expansion. This theoretical prediction had some observational support, but the development of better instruments allowed this prediction to be tested more thoroughly. In 1998, researchers measured the expansion rate of the Universe more carefully. The Hubble Space Telescope observations of very distant supernovae showed that these objects looks fainter than predicted. Therefore, the Universe was expanding slower then than it is today; so this expansion had not been slowing down as thought - it has been accelerating. This was one of the most surprising discoveries in the history of science - and scientists had to explain this phenomena. The only reasonable explanation is that this expansion is accelerating due to a negative pressure caused
by a repulsion force.

Theorists came up with three different explanations for this mysterious force. One, it maybe the cosmological constant that was introduced by Einstein in the theory of General Relativity. Second, there might be some weird energy-fluid that filling the space. Third, there is something fundamentally wrong with General Relativity and gravitational force behaves differently on cosmological scales. Theorists still do not know what is the actual reason of this cosmic acceleration so they called this negative pressure or repulsion force, the dark energy. Very little is known about the dark energy. The only thing we know about it is its responsibility for the expansion of the Universe. Other than that, it is a mystery - and an important one. According to recent observations, this dark energy roughly makes up about 68% of the Universe. The other two components are the dark matter which makes up about 27%, and the rest - everything observed with the limits of our observations, normal matter - adds less than 5%.

Theorists working furiously to try to find out the reason behind this dark energy. One possible explanation for this mysterious dark energy is that, it is an intrinsic property of space. It turns out that space is not empty as it appears has nothing. One property of space is, it is possible for more space to come into existence. Another property of space according to the theory of relativity with the cosmological constant predicts that: “empty space” can possess its own energy. This energy is a fundamental property of space itself - so this energy would not be diluted as space expands. As more of the space comes into existence, more of this energy would appear. Therefore, this particular type of energy would be responsible for making the expansion of the Universe faster and faster. The cosmological constant is a mystery on its own as no one understands why it should even be one of the components of the Universe. Additionally, no one understands why it has the exact right value to cause the expansion to be accelerating.

Another alternate explanation for how space possesses energy comes from the quantum theory. Surprisingly, “the vacuum” is actually not an empty - it is full of
“virtual” particles that form and disappear consistently. Physicists calculated how much an energy is necessary for that empty space, but the answer was disastrous. The number came out to be 10120 times off [1]. That is, a 1 with 120 zeros after it! So, this explanation was ruled out from the possible candidates of dark energy. One more possibility for dark energy is: a new hypothesized field that fills all the space. The effects of that field on the expansion is quite the opposite of that of the normal matter and energy. Theorists have called this field “quintessence”, adopted from the Greek philosophers’ fifth element. Again, there is nothing known about this quintessence - what it is, what it looks like, how and with what it interacts, or what is the fundamental reason of its existence. Thus, it did not solve the mystery of dark energy.

The last possible explanation is that the theory of gravity is not correct on cosmological scales - that is, gravitational force behaves differently from a cosmological perspective in comparison to the traditional view. If this speculation is true, a new theory of gravity would be necessary to explain the expansion of the Universe, the matter clustering, and the behavior of galaxies and galaxy clusters. If indeed it turns out that a new theory of gravity is the answer to the dark energy problem, would it accurately describe the dynamics of the Solar System as general relativity does? Would that theory provide an accurate predictions for our Universe? Physicists have proposed some candidate theories; none of them are compelling though. So, in conclusion, the dark energy mystery continues. Out of all possible scenarios, most recent observations favors the cosmological constant explanation. However, in order to conclusively decide between all these possible candidates of dark energy, we need more observations and more work.

1.3 Statistics of Matter Clustering on Large Scales

One fundamental goal of modern cosmology is to study the matter and energy distribution in the Universe. Galaxy surveys are the essential tools that are used to study
the distribution of matter. The distribution can fix basic cosmological parameters and narrow or even eliminate cosmologies. In order to properly analyze the massive data sets produced from these surveys, cosmologists developed different statistical measures. Any galaxy survey will cover a particular “region” of the Universe, or more technically, a “redshift range.” Regardless of the survey design or the redshift covered, there is a handful of basic techniques that are commonly used to analyze the data sets. According to the recent observations, an underlying dynamics governs how the matter is distributed; in other words, the matter is not just randomly distributed. In our study of the large scale structure, we study how matter and energy are clustered. We find that this clustering is different than a random clustering. To describe this quantitatively, cosmologists developed different statistical measures such as the power spectrum, and the 2-point correlation function. Before discussing these two measures, we must introduce firstly what is known as the density field.

1.3.1 Density Field

Let $\rho(\vec{r})$ the number density of objects at location $\vec{r}$. The quantity $\bar{\rho} = \langle \rho(\vec{r}) \rangle$ is the mean density of the distribution. Define the dimensionless parameter $\delta(\vec{r})$ as

$$\delta(\vec{r}) \equiv \frac{\rho(\vec{r}) - \bar{\rho}}{\bar{\rho}}$$

(1.1)

where $\delta(\vec{r})$ is a measure of the difference of the density at some given point from the average density. This quantity is what cosmologists use to study the evolution of large scale structure. On sufficiently large scales, this quantity has a Gaussian distribution [3]. To interpret this, the density perturbations satisfy Gaussian initial conditions, and the probability of the perturbation amplitude has a Gaussian shape about the mean value.

1.3.2 Two Point Correlation Function

The standard measure of clustering is the two-point correlation function. Generally, the two-point correlation function provides a statistical relation between two vari-
ables located at two different points in space or time in the Universe. As applied to cosmology, the two-point correlation function is essentially a description of the distribution of matter in the Universe. It is a function of distance and calculates the probability of the separation between two galaxies. To describe the correlation function quantitatively, let’s consider the following scenario. Suppose that we have two small regions in the Universe, \( \delta V_1 \) and \( \delta V_2 \), separated by a distance \( r \). Then the expected number of pairs of galaxies with one galaxy in \( \delta V_1 \) and the other in \( \delta V_2 \) is given by \[ \langle n_{\text{pair}} \rangle = \pi^2 \left[ 1 + \xi(r) \right] \delta V_1 \delta V_2 \] (1.2)

In the above equation, \( \pi \) is the average number of galaxies per unit volume. In the above equation, the quantity \( \xi(r) \) can take any value. If \( \xi(r) = 0 \), then the distribution is Poisson; if \( \xi(r) < 0 \), then the distribution is under clustered compared to Poisson; if \( \xi(r) > 0 \), then the distribution is over clustered. The correlation function \( \xi(r) \) can be written as the product of the overdensity function given by equation 1.1 shifted by a distance \( \vec{r} \). Exploiting the cosmological principle we can make the two powerful arguments. Under the assumption of homogeneity, the function is translationally invariant; so it depends only on the vector \( \vec{r} \) separating the two vectors. By a similar way, under the assumption of isotropy, the function is rotationally invariant; so, it depends only on the distance between the two vectors [6]. We can describe this quantitatively as

\[ \xi(\vec{r}) = \langle \delta(\vec{x} + \vec{r}) \delta(\vec{x}) \rangle \] (1.3)

So,

\[ \xi(\vec{r}) = \xi(r) = \langle \delta(\vec{x} + \vec{r}) \delta(\vec{x}) \rangle \] (1.4)

By substituting equation 1.1 in the above equation we get

\[ \bar{\rho}^2 [1 + \xi(r)] = \langle \rho(\vec{x} + \vec{r}) \rho(\vec{x}) \rangle \] (1.5)

For the two small regions \( dV_1 \) and \( dV_2 \) separated by the vectors \( \vec{r}_1 \) and \( \vec{r}_2 \), the expected number of pairs of galaxies, \( n_p \), is

\[ n_p = \bar{\rho}^2 [1 + \xi(r)] dV_1 dV_2 \] (1.6)
In case of a random distribution, the probability of finding a galaxy in the region of \( dV_2 \) is independent of the probability of finding an object in \( dV_1 \). So the expected number of pairs of galaxies in such a random distribution is given by

\[
n_{p,\text{rand}} = \rho^2 dV_1 dV_2 \quad (1.7)
\]

The above equation can be re-written as

\[
n_p = n_{p,\text{rand}} + \rho^2 \xi(r) dV_1 dV_2 \quad (1.8)
\]

Percival gave the following excellent analog to describe the two point correlation function. The following excerpt is quoted from [3] “Imagine throwing down a large number of randomly placed, equal length \( r \), sticks within a survey. Then the correlation function for that separation \( r \) is the excessive fraction of sticks with a galaxy close to both ends, compared with sticks that have randomly chosen points within the survey window close to both ends.” This correlation function is known for its computational challenges. However, it forms a Fourier transform pair with the power spectrum as we will see next.

### 1.3.3 Power Spectrum

The simplest definition of a spectrum is the following: the spectrum of any physical phenomenon is unique; it is a relationship between the magnitude of some parameter against the frequency. This relation is usually represented by a plot of that magnitude against frequency. The field of galaxy surveying has been booming in the last few decades and cosmologists had to introduce new statistical measures to analyze the massive data sets of these surveys. The power spectrum in the case of galaxy surveys shows how much galaxy clustering is happening and at which scales it is occurring. The power spectrum is the basic tool in studying large scale clustering in cosmology [1]. Quantitatively, the power spectrum gives the variance as a function of scale or frequency, and it is the Fourier transform of the correlation function. It is computationally inexpensive. Although different authors are using different
formulas of the power spectrum, the following is the most general expression from
[1]
\[ P(k) = \langle |\delta_k|^2 \rangle \] (1.9)
where \( \delta_k \) are individual Fourier components obtained from the density fluctuations \( \delta(\vec{r}) \)
\[ \delta(\vec{r}) = \frac{V}{(2\pi)^3} \int \delta_k e^{i\vec{k} \cdot \vec{r}} d^3 k \] (1.10)
The correlation function and the power spectrum form a Fourier pair
\[ P(k) = \int \xi(r) e^{ik \cdot r} \, d^3 r \] (1.11)
\[ \xi(r) = \int P(k) e^{-ik \cdot r} \frac{d^3 k}{(2\pi)^3} \] (1.12)
In chapter 3, we will discuss in detail how we calculated the power spectrum.

1.4 Fractals in Cosmology

In this section we introduce the basic theoretical background of fractals necessary to our project, and to provide the reader that may not be familiar with fractals their basic machinery. The subject of fractal geometry has an extensive literature, so we are not going to provide a comprehensive discussion here. The interested reader is referred to Mandelbrot [13] to know about the original motivations behind fractals. For the mathematically inclined reader, Falconer’s book [11] is excellent.

1.4.1 On The Meaning of Being A Fractal

The name fractal was introduced by the late Benoît B. Mandelbrot to describe geometrical figures which are irregular, not smooth, and repeats itself over different scales [8]. Mandelbrot’s seminal work in the 70s and 80s culminated in his book (The Fractal Geometry of Nature) provided the foundations of the field of fractal geometry. Inspired by his work, the field of fractal geometry saw an explosive growth
over the next two decades, and the fractal analysis of data sets has established itself as a valuable tool in the physical sciences, medicine, geography and geology, and finance. The term fractal has its roots in the Latin adjective “fractus” which means “to break” or “to create irregular fragments.” Later, Mandelbrot offered the definition: a fractal is a shape made of parts similar to the whole in some way. Falconer [11] give the following rigorous description. A set, $F$, is said to be a fractal if:

1. $F$ has a fine structure, i.e., detail on arbitrarily small scales.

2. $F$ is too irregular to be described in traditional geometrical language, on both small and large scales.

3. $F$ has some form of self-similarity, i.e., the structure repeats itself over different scales, either approximately or statistically.

4. The Fractal Dimension $D$ of $F$ is interconnected with its topological dimension $m$. $D$ is either greater than, equal to, or less than $m$.

Mandelbrot’s and Falconer’s definitions are different to some extent. However, both of them agree that fractals are structures that are self-similar in some way and they take fractional dimensional values. Indeed, there is a consensus among specialists to characterize fractals, at least, as self affine processes. Falconer’s definition give the flexibility to describe shapes as fractals if one or more of the properties above satisfied. That is, an object is said to be a fractal if the properties of that object can be described by the above mentioned definitions. There is no standard definition for fractals; indeed, Falconer provided his view on this issue by saying: “The definition of a fractal should be regarded in the same way as the biologist’s definition of life. There is no hard definition, but a list of properties that has to be satisfied.” For example, a mathematician might define an object as a fractal if the object is extremely irregular to be described by traditional geometry. On the other hand, a cosmologist would refer to galaxies and galaxy clusters as fractal objects because
they exhibit some sort of self-similarity. Both the definitions of the mathematician and the cosmologist are true and valid, on the basis of Falconer’s criterion.

The Fractal Dimension

The fundamental goal of fractal analysis is to calculate the fractal dimension $D$ of the data set under consideration. The definitions of fractal dimension are myriad, but the most general definition is: a dimension $d$, is a description of the capacity of a set $F$, to fill the space. That is, it is a measure of the degree of the irregularities present in that set as viewed at different scales. Another way to think about this is that, this dimension is an index of the prevailing complexity of the set [8]. As the name implies, the fractal dimension need not be an integer number - indeed it is in most case not an integer. Different authors use different ways to estimate the fractal dimension, and we are going to present our method in chapter 3.

1.4.2 Examples of Fractals

Fractals are ubiquitous in nature - they are found everywhere in the Universe from the smallest to the largest scales and from natural to artificial structures. One classical problem of fractal geometry is the length of coastlines; coastlines are examples of natural fractal structure. Mandelbrot provided a wonderful discussion on this problem and its solutions. It turns out that determining the length of the coastline of a country is not a trivial - it is not simple as it first appears. Indeed, the measured length depends on the scale of measurement; the smaller the increments of measurements, the longer the measured length of the coastline. By plotting the length of the used scale in measurements versus the measured length of the coastline on a log-log plot, the relation is a straight line. The slope of this line is the fractal dimension of the coastline which lies between $(d = 1 - 2)$. Another recognized fractal object that is artificial is the Koch Curve illustrated in figure 1.4. It is a mathematical fractal object that was first described by the mathematician Helge Von Koch in 1904. To
build it, start with an equilateral triangle, remove the inner third of each side, build another equilateral triangle instead of the removed side, and repeat this steps over and over again. The Koch Curve or snowflake is, of course, not a line; its dimension is not 1. It also does not fill a plane, so its dimension is not 2. Its dimension lies between \( (d = 1 - 2) \). This object satisfies all the properties given by Falconer - it has a fine structure, exhibits self similarity, and too irregular. As such, the Koch Curve is one of the best objects that represent fractals.

1.4.3 Fractal Analysis of Galaxies

The power spectrum and the two point correlation function are established statistical measures to classify large scale structure since the earliest days of galaxy surveys. Following the development fractal geometry, fractal analysis of galaxies has been
added to the statistical measures of cosmology to better understand the large scale structure. This kind of analysis has become a tool for studying matter clustering in the Universe. The goal of fractal analysis of galaxy distributions is to test whether the distribution under study has fractal properties or not, by calculating the fractal dimension of the distribution, $D$.

From a cosmological perspective as applied to galaxy distribution, the fractal dimension is a measure of the degree of inhomogeneity in the distribution, i.e., how irregular is the distribution (departure of the distribution from regularity). It provides information about the dominance of voids in structure, hence, it is an important tool in studying matter clustering. Depending on the relation between the fractal dimension $D$, and the corresponding topological dimension $n$ where the system is embedded, the degree of homogeneity and inhomogeneity of structure can be obtained. If the value of $D$ is smaller than the value of $n$, the distribution is said to be irregular. Systems like galaxies that assumed to live in three dimensional topology are said to be regular if $D = n = 3$. In such a case, the distribution is homogenous. On the other hand if $D < 3$, the galaxy systems are irregular, and decreasing values of $D$ implies an increasing inhomogeneity or irregularity [9]. Systems that are described by one fractal dimension are said to be single fractals, and they are the simplest possible fractal structures in nature. More complex structures or distributions are likely to have different values of the fractal dimension over different scales or distance ranges. That is, the fractal dimension $D$ is a function of the distance $d$, so $D = D(d)$. In that case, the fractal system under consideration is said to be multifractal [13].

**Previous Results**

Fractal analysis of galaxies has gained an increasing popularity among cosmologists since the 80s and the literature has several studies on that work. The earliest cosmological model that described galaxies as fractal systems was the hierarchical model of Wertz (1970, 1971) [8]. Unfortunately, the ideas of fractal geometry had
not been developed then, so his model remained largely ignored. Peebles (1980, 1993) and Davis (1997) developed their model using the traditional galaxy clustering view. Davies and Peebles reported $D = 1.2$ using the correlation function at 5 Mpc. Pietronero, Montuori, and Sylos Labini [7] made probably the most comprehensive analysis in the 90s. They used catalogs from the CfA, Perseus-Pisces, SSRS, IRAS, Stromlo-APM, LEDA, Las Campanas and ESP for galaxies and Abell and ACO for clusters. Their method was based on its entirety on the concepts of modern statistical physics - not on the traditional clustering tools. Their conclusions were quite remarkable: galaxy distributions are highly irregular, inhomogenous, and self-similar. Their most tantalizing yet surprising claim was that the distribution of visible matter in the universe is fractal and not homogeneous! Their claim rejuvenated the research on the observational homogeneity of (see[12]). Table 1.1 contains the results of their analyses and it is reproduced from their original 1997 paper. Most recently, Saavedraa, Iribarrema, and Ribeiro reported $D = 0.5$ and $D = 1.4$ using the galaxy volume number densities [9].

<table>
<thead>
<tr>
<th>Galaxy Survey</th>
<th>The Measured Fractal Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>CfA1</td>
<td>$1.7 \pm 0.2$</td>
</tr>
<tr>
<td>CfA2</td>
<td>$2.0$</td>
</tr>
<tr>
<td>PP</td>
<td>$2.0 \pm 0.1$</td>
</tr>
<tr>
<td>SSRS2</td>
<td>$2.0$</td>
</tr>
<tr>
<td>LEDA</td>
<td>$2.1 \pm 0.2$</td>
</tr>
<tr>
<td>LCRS</td>
<td>$1.8 \pm 0.2$</td>
</tr>
<tr>
<td>IRAS</td>
<td>$2.0 \pm 0.1$</td>
</tr>
<tr>
<td>ESP</td>
<td>$1.9 \pm 0.2$</td>
</tr>
</tbody>
</table>

Table 1.1: Results of the fractal Analysis of different surveys from the 1997 Pietronero, Montuori and Labini paper [7]. $D$ is the fractal dimension.

1.5 The Sloan Digital Sky Survey SDSS

To study the large scale structure, we need data from large surveys that represents the matter and energy distribution in the Universe. Observational cosmology has
been successful in designing different surveys that allowed cosmologists to precisely calculate important cosmological parameters in the last few decades. In this thesis, we will be using data from the Sloan Digital Sky Survey (SDSS). The SDSS is one of the most ambitious cosmological survey that has ever been designed, with a goal to produce the most detailed 3-d map of the Universe. The SDSS is an all sky spectroscopic survey that has been observing the Universe using the 2.5 m Sloan Foundation Telescope at Apache Point Observatory (APO), NM, USA since April 2000. The SDSS operates different sub-surveys that targeting different projects that covers wide fields of research. Some of the projects are the (APOGEE), the extra-solar planet discovery (MARVELS), galactic structure (SEGUE-2), and the Baryonic Oscillations Spectroscopic Survey (BOSS) which we used. The survey covers spectroscopic analysis of target-selected objects in the night sky covering approximately 1/3 of the celestial sphere. As we write this thesis, SDSS is undergoing its fourth major survey, SDSS-IV using an additional 2.5 meter telescope at Las Campanas Observatory in Chile. The previous three surveys SDSS-I, SDSS-II, SDSS-III were conducted between 2000-2014. The data has been made available to the public for scientific research through different data releases (known as DR). The data release include both the raw spectroscopic data and object catalogues [26, 27].

1.5.1 Baryonic Oscillations Spectroscopic Survey BOSS

In this thesis we used data release DR 13 from the the Baryon Oscillation Spectroscopic Survey (BOSS), and the extension of this project for SDSS-IV, eBOSS. The redshift covered in these surveys BOSS/eBOSS is $0 < z < 0.8$ as seen in figure 1.6. The surveys covers two continuous regions of sky covering approximately 10400 deg$^2$, with a total of 2.5 million spectra [27]. Once the data were obtained from the target objects, the spectra were processed by the BOSS data pipeline to classify the objects types into (star, galaxy, quasar, etc.). After that, the classified objects are grouped into different catalogs and different research groups use the catalogs, they want according to their preferences.
Figure 1.5: The SDSS coverage. The BOSS survey covers the redshift range $0 < z < 0.8$. Source: reproduced from the SDSS website.

1.5.2 Mock Galaxy Catalogs

In addition the BOSS/eBOSS surveys, we used also the Mock Galaxy Catalogs produced by the SDSS. Mock Catalogs are used to optimize the survey, to create the covariance matrix for the clustering measurements, and to test the accuracy of the analysis. The technique to create the mock catalogs is the N-body simulation. The mock catalogs we used are the SDSS, and they were created using the quick particle mesh method with the benchmark cosmological model as outlined in [25]. The model is flat $\Lambda$ CDM model with $\Omega_m = 0.274$, $\Omega_b = 0.046$, $\Omega_\Lambda = 0.726$, $H = 70$
km s$^{-1}$ Mpc, $n = 0.95$, and $\sigma_8 = 0.8$. The assumptions made in creating the mock catalogs are: (1) Galaxies form and remain in the potential wells of the dark matter halos, (2) Halos represent overdense regions of the mass field that is (100-300) times greater than the mean density that arise form the non-linear gravitational collapse. The steps of creating the mock catalogs are:

1. Predicting the evolution of the mass field.
2. Locating the dark matter halos and characterize its properties.
3. Populating the halos with mock galaxies.
4. Apply survey characteristics to the galaxies box.

An updated version of the TreePM$^2$ code was used to evolve (3000)$^3$ particles with mass a of $5.9 \times 10^{10} h^{-1} M$ in a box of side $(2750)h^{-1}$ Mpc. The number of data points we used was approximately 32 million. The initial conditions of this simulation were: (1) Displacing the particles from a regular grid using 2nd order Lagrangian Perturbation theory, (2) $z = 25$, (3) The RMS displacement is 1.4 Mpc, (4) Time steps are sampled using the 2nd order leap frog method (friends of friends algorithm), (5) In populating galaxies by evolving the density field, the particles are selected such that they match the one and two point statistics of the dark matter halos.

1.6 Goals and Motivations

As this point, it must be clear what we are trying to accomplish in this project. Our goal is to study the matter and energy clustering in the Universe using the fractal analysis approach. Can galaxies, clusters of galaxies, and the large scale structure of the Universe be described as fractal systems or not? Are the galaxies distributions homogenous or inhomogenous? What dominates the large scale structure of the Universe? These questions are what motivates our research, and we will present the answers to them by the end of chapter 5. To do this, we developed an algorithm
to estimate the fractal dimension of galaxies from the SDSS catalogs. There are
different methods to estimate the fractal dimension of galaxy catalogs, and different
authors have reported different values of $D$ according to their methodologies. In
this project, we developed an algorithm to calculate the angular fractal dimension
of galaxy catalogs from the BOSS survey and the Mock Catalogs of the SDSS as
a function of the cosmological comoving space $D = D(r)$. We will use the BOSS
galaxies and the Mock Catalogs of the SDSS to test the validity of our algorithm.
Unlike other authors that had developed their algorithms based on traditional sta-
tistical measures, our algorithm is based primarily on the wavelet transform. We
will extensively discuss the wavelet transform in the next chapter.
CHAPTER 2

Wavelets

“For whosoever hath, to him shall be given, and he shall have abundance: but whosoever hath not, from him shall be taken away even that which he hath”
Matthew 13:12

In this chapter we will discuss wavelets, the mathematical tools of our research. In section 2.1, we will introduce wavelet theory and describe what wavelets are and explore their machinery. After that, we will introduce a specific type of wavelets we used in our research known as Daubechies1, or Haar wavelets. In section 2.2, we will describe two particular wavelet transforms known as: Discrete Wavelet Transform (DWT) and Wavelet Packet Transform (WPT). We will discuss how to apply the wavelet transform to one-dimensional data. After that in section 2.3, we will extend the transformation to the two dimensional case of the galaxies. We will conclude the chapter in section 2.4 by presenting the advantages and disadvantages of wavelet transform and compare it with the Fourier transform.

2.1 Definition and Properties

The simplest definition of wavelets is they are mathematical functions used to divide a given signal into different components. Wavelets are a tool used in signal analysis that are in some way similar to the Fourier transform. The main difference is that the wavelet transform is localized in time (or space) as well as the frequency (or scale) domain, while the Fourier transform is localized in the frequency domain. The result of a Fourier analysis shows which frequencies have occurred in the signal,
but information about where in the signal they have occurred is difficult to obtain since this information is located in the phases. A wavelet generally consists of two functions, a low pass and a high pass filter. In wavelet transform, the signal is decomposed into two sub-signals. One of these is a localized density which is carried by the so-called approximation coefficients and the other is a localized fluctuation carried by the detail coefficients. There are different kinds of wavelets like the Haar, Daubechies, Mexican Hat, and others. In this thesis work, we used the Daubechies 1 or Haar wavelets. We will begin by describing the Haar wavelet because it is the simplest of the Daubechies wavelets. We will use it to explain wavelets conceptually and mathematically. The important mathematical properties of wavelets can be summarized as follows: they integrate to zero over infinity; they are well localized; and they exhibit oscillatory behavior.

2.1.1 Haar Wavelet

As we said earlier, wavelets consist of low pass and high pass filter. The step function is perhaps the simplest function that represents a low pass filter and is the key to building both the Haar scaling function and wavelet. Mathematically, begin with a unit step function

\[
\phi(x) = \begin{cases} 
1 & 0 \leq x < 1 \\
0 & \text{otherwise}
\end{cases} 
\]  

(2.1)

This step function is equal to 1 if \( x \) lies in the interval \([0, 1)\) and zero in any other interval as seen in figure 2.1. As we will see later, this step function is used to build the low pass filters. For any interval \( u, w \), equation (2.1) can be written as

\[
\phi_{[a,u]}(r) = \phi\left(\frac{r - u}{w - u}\right) \begin{cases} 
1 & u \leq r < w \\
0 & \text{otherwise}
\end{cases}
\]  

(2.2)

At this point we can define the Basic Haar scaling function which is built from the combination of two step functions

\[
\phi_{[a,b]} = \phi_{[a,\frac{a+b}{2}]} + \phi_{[\frac{a+b}{2},b]}
\]  

(2.3)
In other words, the scaling function breaks up an interval \([a, b]\) into two step functions split evenly across the interval. How this is used to analyze signals is presented later in this chapter.

![Haar Scaling function](image)

Figure 2.1: The Haar Scaling function \(\phi(r)\) along the interval \((0, 1)\). The function is equal to zero outside the interval \([0, 1)\). Source: reproduced from Texas A and M University Math 414 - Spring 2004 Course Website.

Similar to the step function, we can introduce another function called the wavelet function \(\psi\). This wavelet function acts as the high pass filter as we will see later in this chapter. The Haar wavelet is built also out of two step functions as

\[
\psi_{[a,b]} = \phi_{[a, \frac{a+b}{2}]} - \phi_{[\frac{a+b}{2}, b]} \quad (2.4)
\]

In other words, the wavelet function breaks up an interval \([a, b]\) into two step functions split evenly across the interval as seen in figure 2.2. How this wavelet is used to analyze signals is presented later in this chapter. We can write the wavelet over a generic interval as

\[
\psi_{[u,w]}(r) = \begin{cases} 
1 & 0 \leq r < \frac{1}{2} \\
-1 & \frac{1}{2} \leq r < 1 \\
0 & \text{otherwise}
\end{cases} \quad (2.5)
\]
The scaling function \( \phi \) and the wavelet \( \psi \) are the basic building blocks of the wavelet transforms [14]. By using the scaling and the wavelet functions, one can represent the signal as a linear combination of \( \phi \) and \( \psi \) as we will show later. The analog to this representation is the use of the sin and cos functions to expand the signal in the Fourier methods. It became apparent that both the scaling function and the wavelet have some physical meaning; they represent local densities and local fluctuations in the signal, respectively. In other words, they represent the local approximations and local differences, in the density field.

![Figure 2.2: The Haar wavelet \( \psi(r) \) along the interval (0,1). Source: reproduced from Wikimedia Commons, the free media repository.](image)

### 2.2 Wavelet Transforms

Having discussed how wavelets are constructed, we now introduce the wavelet transform. In particular, we will introduce the Discrete Wavelet Transform (DWT). We will extend the discussion later to the other type of the wavelet transform we used in this thesis, the Wavelet Packet Transform (WPT). In order to understand the WPT, it is first necessary to illustrate the DWT. We do so by introducing briefly an explanation of the signal approximation using wavelets.
2.2.1 The Discrete Wavelet Transform

Haar DWT

In this section, we will explain how we can expand or represent our original function in terms of the two basis functions we introduced, the scaling function and the wavelet, $\phi$ and $\psi$ respectively. These two functions represent again the local densities and the local fluctuations of the signal. To do that, we first write the experimental discrete signal as a vector [14]

\[ s = (s_0, s_1, ..., s_k) \quad (2.6) \]

where $s_0$ is the first sample point of the signal and the samples are taken evenly and $j$ is some integer. The next step is approximating the signal by writing it in terms of the scale function $\phi$ as

\[ f(x) = \sum_{j=1}^{2^N-1} s_j \phi_{[x_j, x_{j+1})}(x) \quad (2.7) \]

Now, we will replace every successive pair of points with a wider scaling function and a wavelet. We will expand the function in terms of a new set of basis functions. These new basis functions are the $\phi$ and $\psi$ functions. We can represent the new basis functions mathematically as

\[ \phi_{(a,b)} = (x) = \phi(x)_{(a, \frac{a+b}{2})} + \phi(x)_{(\frac{a+b}{2}, b)} \quad (2.8) \]

\[ \psi_{(a,b)} = (x) = \phi(x)_{(a, \frac{a+b}{2})} - \phi(x)_{(\frac{a+b}{2}, b)} \quad (2.9) \]

We now introduce two new definitions: the approximation coefficients $a_j$ and the detail coefficients $d_j$. The approximation coefficients are the coefficients of the wider step functions, and the detail coefficients are the coefficients of the wavelet functions. They are represented mathematically as

\[ a_j = \frac{s_j + s_{(j+1)}}{2} \quad (2.10) \]
\[ d_j = \frac{s_j - s_{j-1}}{2} \quad (2.11) \]

We can realize that the approximation coefficients are the average of the signal at two successive points, and the detail coefficients are the difference of the two points. So we can write the approximation of the function in terms of the new coefficients \( a \), and \( d \) as

\[
f(x) = \sum_{j=1}^{2^{N-1}} s_j \phi(x_j, x_{j+1})(x) = \sum_{j=1}^{2^{(N-1)-1}} a_j \phi(x_j, x_{j+1})(x) + \sum_{j=1}^{2^{(N-1)-1}} d_j \psi(x_j, x_{j+1})(x) \quad (2.12)
\]

In the above equation, we have written our function as an expansion of the approximation and detail coefficients \( a_j \), and \( d_j \), and the basis functions \( \phi \), and \( \psi \). The fundamental idea of the wavelet transform is we act two filters (low and high pass) on the signal to separate the components of the signal according to the frequency they represent as seen in figure 2.3.

Figure 2.3: The act of the low and high pass filters on a simulated signal of a sin wave with noise.
We now introduce an example that demonstrates how this transform works. Consider the signal

\[ \vec{s} = (2, 4, 6, 2, 1, 7, 8, 2) \] (2.13)

Using the approximation equation 2.7, the signal can be written as

\[ f(x) = 2\phi_{0,1} + 4\phi_{1,2} + 6\phi_{2,3} + 2\phi_{3,4} + 1\phi_{4,5} + 7\phi_{5,6} + 8\phi_{6,7} + 2\phi_{7,8} \] (2.14)

The first pass of the wavelet transform will transform the signal into a pair of approximation and detail coefficients. The approximation coefficients are labeled by \( a_j \) and the detail coefficients are labeled by \( d_j \). So we get the first pair of those coefficients as

\[ a_1 = \frac{s_0 + s_1}{2} = \frac{2 + 4}{2} = 3 \] (2.15)

\[ d_1 = \frac{s_0 - s_1}{2} = \frac{2 - 4}{2} = -1 \] (2.16)

Similarly, we get the following pairs of \( a_j \) and \( d_j \) for the remaining signal components

\[ a_2 = 4 \quad d_2 = 2 \] (2.17)

\[ a_3 = 7 \quad d_3 = -3 \] (2.18)

\[ a_4 = 5 \quad d_4 = 3 \] (2.19)

We can write the signal as a combination of the \( a_j \) and \( d_j \) coefficients as

\[ f(x) = 3\phi_{0,2} + 4\phi_{2,4} + 4\phi_{4,6} + 5\phi_{6,8} - 1\psi_{0,2} + 2\psi_{2,4} - 3\psi_{4,6} + 3\psi_{6,8} \] (2.20)

The first four terms in the above equation corresponds to the the approximation coefficients, and the latter four terms corresponds to the detail coefficients of the signal. Again, the approximation coefficients represents the average between two successive points, and the detail coefficients represents the difference between the
same two points. We can write the signal vector after the first pass of the wavelet transform as
\[ \vec{s} = (3, 4, 4, 5, -1, 2, -3, 3) \] (2.21)
where we grouped the approximation coefficients together in the first four terms, and the detail coefficients together in the latter four terms. This process concludes the first pass of the Discrete Wavelet Transform DWT. Of course, in virtually all the signals we have, we will need to apply another pass of the DWT. To do that, we repeat the process illustrated above again on the transformed signal vector given by the above equation. The main difference, however, is that we perform this transform only on the approximation coefficients. We do not apply the transformation anymore on the detail coefficients. So the last four numbers in the transformed signal vector will remain the same. We will denote the coefficients obtained from the second pass by the notation \( a_{j}^{n+1} \), and \( d_{j}^{n+1} \), where \( n \) is the number of the pass. The first four terms are transformed to
\[ a_{1}^{2} = \frac{3 + 4}{2} = 3.5 \] (2.22)
\[ d_{1}^{2} = \frac{3 - 4}{2} = -\frac{1}{2} \] (2.23)
\[ a_{2}^{2} = 4.5 \quad d_{2}^{2} = \frac{1}{2} \] (2.24)
And we can re-write the signal approximation in exactly the same way as we wrote it earlier. So the new transformed signal after the second pass is
\[ \tilde{s} = (3.5, -\frac{1}{2}, 4.5, \frac{1}{2}, -1, 2, -3, 3) \] (2.25)
At this point, we can only apply one more pass to the transformed signal above following the same procedures. We get the final signal vector
\[ \tilde{s} = (1.5, 2, 2.5, 2, -1, 0, 4, 1) \] (2.26)
We can observe that the four terms \((-1, 0, 4, 1)\) have been carried out unchanged
after the first pass because we don’t apply the transform on them anymore as seen in figure 2.4.

\[ s = (s_0, s_1, s_2, s_3, s_4, s_5, s_6, s_7) \]

\[ a_0, a_1, a_2, a_3 \]

\[ d_0, d_1, d_2, d_3 \]

\[ a_0, a_1 \]

\[ d_0, d_1 \]

\[ a_0 \]

\[ d_0 \]

Figure 2.4: The DWT tree. The original signal in the top is transformed into the approximation and detail coefficients after the first pass or sweep \( j = 1 \). Then at level \( j = 2 \), we perform the DWT again only on the approximation coefficients, and the detail coefficients are left without any more transformation. We proceed further with more passes or levels following the same procedures.

We can realize some interesting properties of the Haar DWT. First, we can observe is that after each pass of the DWT, the number of the approximation coefficients is reduced by two. However, the number of the detail coefficients remains the same throughout the transformation because we don’t operate on them anymore after the first pass as seen in figure 2.4. The second property we can observe is that there is no information lost on this transformation, in other words, the transformation is reversible. We can get the original signal form the transformed signal by inverting the transform. Mathematically, we represent this process by

\[ s_j = 2(a_j + d_j) \quad s_{j+1} = 2(a_j - d_j) \]  \hspace{1cm} (2.27)

So we can write our signal as

\[ \bar{s} = (a_0, a_1, \ldots; d_0, d_1, \ldots) \]  \hspace{1cm} (2.28)
2.2.2 The Wavelet Packet Transform

The Wavelet Packet Transform (WPT) is similar to the DWT. In both transforms, the method for acquiring approximation and detail coefficients remains the same, as explained in the DWT section. There is one main difference between the WPT and the DWT. While in the DWT we perform the transform after the first pass on the approximation coefficients only, in the WPT we perform the transform on both approximation and detail coefficients. In comparing the DWT with the WPT, the major difference is that the WPT has better frequency resolution than the DWT. As a result, we end up with more information about our signal in the frequency domain [21]. To distinguish between the approximation and the detail coefficients at various levels, we will introduce the following notation. As usual, we denote the approximation coefficients by $a_j$ and the detail coefficients by $d_j$.

After we perform the first pass on the signal vector, we get the approximation and detail coefficients of the original signal as denoted, for example, by $a_0, a_1, a_2, a_n$ and $d_0, d_1, d_2, d_n$, respectively. In the wavelet packet, we pass both branches of the transformed signal through the high pass and low pass filters. The result on the approximation coefficients, $a_0, a_1, a_2, a_n$, is two new approximation coefficients, and two new detail coefficients spanning the same total frequency range, but now divided into narrower frequency bands. We denote those new approximation coefficients by $aa_0$, and $aa_1$ and new detail coefficients by $ad_0$, and $ad_1$. The coefficients $aa_0$ are the approximation coefficients obtained by acting on the $a_i$ approximation coefficients. The coefficients $ad_1$ are the detail coefficients obtained by acting on the $a_i$ approximation coefficients.

The same applies when we act on the detail coefficients $d_0, d_1, d_2, d_n$. We will get two new approximation coefficients $da_0$, and $da_1$, and two new detail coefficients $dd_0$, and $dd_1$. The coefficients $da_0$ are the approximation coefficients obtained by acting on the detail coefficients $d_i$. The coefficients $dd_1$ are the detail coefficients obtained by acting on the detail coefficients $d_i$. This notation carries out throughout
the entire process of the WPT tree. This process of tracking the approximation and detail coefficients is illustrated in figure 2.5.

As defined previously, the approximation and detail coefficients have some physical meanings; they represent the local densities and local fluctuations in the signal. In the case of the WPT, the repeated action of the low and high pass filters at different levels of decomposition separate the signal into components with different frequencies. We can associate these coefficients with the low and high pass filters. The action of the low pass filter on the signal give the approximation coefficients, and the action of the high pass filter on the signal give the detail coefficients. Consider the coefficients $aa_i$, they contain information about low-frequency band of the original signal. On the other hand, the coefficients $dd_i$ contain information about the high-frequency band of the original signal as seen in figure 2.5. To illustrate how the WPT works, here is an example. We can use the same signal vector we used in the DWT. Recall that our signal vector is

$$\vec{s} = (2, 4, 6, 2, 1, 7, 8, 2)$$

After applying the first pass of the WPT, we get the transformed signal

$$\vec{s} = (3, 4, 4, 5, -1, 2, -3, 3)$$

Applying a second pass to the transformed signal on both the approximation and detail coefficients will give

$$\vec{s} = (\text{aaa}_0, \text{aad}_0, \text{ada}_0, \text{dab}_0, \text{dab}_1, \text{daa}_0, \text{dda}_0, \text{ddda}_0) = (3.5, 4.5, -0.5, 0.5, 0.5, 0, -1.5, -3)$$

Applying a third pass on the above transformed signal will give us the final transformed vector, which composed of completely transformed coefficients.

$$\vec{s} = (\text{aaa}_0, \text{aad}_0, \text{ada}_0, \text{dab}_0, \text{dab}_1, \text{daa}_0, \text{dda}_0, \text{ddda}_0) = (4, -0.5, 0, -0.5, 0.25, 0.25, -2.25, 0.75)$$
Figure 2.5: The WPT tree. The vector $S$ is the original signal. After the first pass or sweep $j = 1$, the original signal vector is transformed into approximation coefficients $a_i$ and detail coefficients $d_i$. The approximation coefficients are the result of the action of the low pass filter on the signal. The detail coefficients are the result of the action of the high pass filter on the signal.

### 2.3 2-D Wavelet Transforms

We have introduced the 1-D DWT and WPT. In our research we have to extend our treatment to the 2-dimensional data. In the case of 1-D data, we had a signal represented by a vector of length $2^j$ where $j$ is an integer. In the 2-D case, our signal is represented by a $2^j \times 2^j$ matrix. To more clearly illustrate, consider the following 2-d signal matrix that represents the sample

$$s = \begin{pmatrix} s_{0,0} & s_{0,1} \\ s_{1,0} & s_{1,1} \end{pmatrix}$$

(2.29)

The 2-d wavelet transform is done in the same way a 2-d Fourier transform is done. That is, we will perform the 1-d transform twice: the first time, on the rows of the matrix, the second time, on the columns of the matrix. The order in which we perform the transform does not matter. We begin by performing the 1-d transform
on the rows firstly
\[
\begin{pmatrix}
\frac{s_{0,0} + s_{0,1}}{2} & \frac{s_{0,0} - s_{0,1}}{2} \\
\frac{s_{1,0} + s_{1,1}}{2} & \frac{s_{1,0} - s_{1,1}}{2}
\end{pmatrix}
\] (2.30)

Then, perform the transform on the columns
\[
\begin{pmatrix}
\frac{(s_{0,0} + s_{0,1}) + (s_{1,0} + s_{1,1})}{4} & \frac{(s_{0,0} - s_{0,1}) + (s_{1,0} - s_{1,1})}{4} \\
\frac{(s_{0,0} + s_{0,1}) - (s_{1,0} + s_{1,1})}{4} & \frac{(s_{0,0} - s_{0,1}) - (s_{1,0} - s_{1,1})}{4}
\end{pmatrix}
\] (2.31)

As usual, we have the approximation coefficients \(a_i\), and the detail coefficients \(d_i\). In addition to those two coefficients, we have an additional two coefficients: the horizontal detail coefficients \(h_i\), and the vertical detail coefficients \(v_i\). As a matrix representation,
\[
\begin{pmatrix}
a_i & h_i \\
v_i & d_i
\end{pmatrix}
\] (2.32)

As it is the case with the 1-d data, we group the coefficients together according to the type of filter (high or low pass) that has acted on the data. In the case of 2-d DWT, we will only perform the transform on the approximation coefficients as is the case with the 1-d DWT. For the WPT, we will perform the transform on all types of coefficients we have in equation 2.32. To track the different coefficients obtained at different passes, we will use the same notation we developed to track the coefficients in the case of the 1-d WPT. To show this, we start as usual with the original signal matrix. Then, we apply the WPT to get the transformed matrix that has 4 different types of coefficients. Then, we apply the transform for another pass on all the coefficients \(a_i, d_i, h_i, v_i\). We will get another 4 sets of coefficients that can be grouped into: 4 \(a_i\) coefficients, 4 \(d_i\) coefficients, 4 \(v_i\) coefficients, and 4 \(h_i\) coefficients. We act on these coefficients with both sets of filter to get the new coefficients now grouped into finer frequency resolution bands as seen in figure 2.6. For example, we can act on the 4 \(a_i\) coefficients and get one coefficient \(aa_{ij}\), one coefficient \(ha_{ij}\), one \(da_{ij}\) coefficient, and finally one \(va_{ij}\) coefficient. Figure 2.6 illustrates the two dimensional WPT.
Figure 2.6: The two dimensional Wavelet Packet Transform Tree. The low and high pass filters acts on all types of coefficients at each pass. Each set of coefficients is represented by a matrix [37].

2.4 Why Wavelets?

As previously mentioned, wavelets are used in signal analysis, they are in some sense similar to the Fourier transform. They are useful for both one and two-dimensional signal in frequency space. We can think of wavelets as a microscope for fine details examination. This process of revealing the details is done by partitioning the signal into different coefficients that have different frequencies and different energies. The wavelet transform gives information about how energy is partitioned on position and scale. In comparing the wavelets transform to the Fourier transform, the wavelet transform is localized in both the time and the frequency domains, but the Fourier transform is localized only in the frequency domain. As a result, the Fourier transform gives energy as a function of frequency in the signal, but it has no information about where (or when) in energy is located. As with the Fourier transform, there is no information lost in wavelet decomposing a signal and the original signal can be faithfully reconstructed by inverse wavelet transforming.

For our work, a key property of wavelets is their self-similarity in both space and scale. Given that the fractal behavior of any physical system is extended over a finite scale, or the fractality is localized, and that fractal behaviour implies self-
similarity, the localization and self-similar properties of wavelets are optimal for use in fractal analysis. However, wavelets do have disadvantages. One major drawback of wavelets is because each pass of the wavelet reduces the resolution by half, the frequency resolution is limited. In addition, with Fourier transforms, the signal is expanded in terms of sin and cos functions which are well understood. On the other hand, wavelet basis functions are often irregular, and have no closed analytic form. To recap, the general recipe for the wavelet transform works as follows

1. Decompose the original signal into a new sub-signal by using the low pass and high pass filters.

2. The signal now is partitioned into two sub-signals, namely the approximation and detail coefficients. We operate the two filters again on those coefficients.

3. Further decomposition is applied on the low-pass filtered signal in the case of the DWT, and on both the low-pass and high-pass filtered signal in the case of the WPT.

**Why Wavelet Packet Transform** The key reason why we chosen the WPT over the ordinary wavelet transform is that it offers a richer signal analysis. In the case of DWT, the information we extract from the signal (i.e., power spectrum) is embedded in the detail coefficients and spread over a relatively wide frequency band. In the case of WPT, the decomposition is applied on both branches of the wavelet tree, so we end up with more frequency bands and therefore, a finer measure of the signal. This can be seen in figures 2.7. Each of the nodes (i.e boxes) has a corresponding frequency band which allows us to extract information from the signal not accessible using the ordinary wavelet. As a result, the WPT offers a better way to examine the signal. The information we extract from the WPT (i.e., power spectrum) is embedded in both the approximation and detail coefficients. From a signal analysis perspective, the Nyquist frequency range in the case of the WPT have more frequency bands than the DWT. Therefore, the WPT offers a finer resolution
to examine the details of the signal. In the next chapter, we will introduce how we used the WPT in our analysis.

Figure 2.7: Comparison between the WPT and DWT. In the case of DWT, one branch of the leaves is terminated after the first pass so the leaves have only one branch. For the WPT, the leaves do not terminate at any resolution. As a result, we end up with more and narrower frequency bands to examine the signal. The sweep number is $j$, and $a$ and $d$ are the approximation and detail coefficients, respectively.
CHAPTER 3

Estimating the Fractal Dimension $D$

“Somethings - a lot of things, are not meant to be known in life.”
Anonymous

In this chapter of the thesis, we will describe in details the algorithm we developed to estimate the Fractal Dimension $D$ of the SDSS galaxies. We begin the chapter by providing an overview of the workflow of our method. In section 3.2, we will describe how we processed the raw SDSS data to pixelize the celestial spheres. In section 3.3, we will explain in detail how we estimated the power spectrum and other auxiliary processes we applied. In section 3.4, we introduce the Hurst Exponent and describe how it is calculated. In the final section of this chapter, we will explain how to calculate the fractal dimension $D$. Whenever appropriate, we present the actual MATLAB codes we used in our calculations. The codes we used were developed primarily by Jesús Pando and Loay Khalifa. Detailed descriptions of the codes are provided in the appendices.

3.1 Workflow

The workflow of our algorithm is outlined in figure 3.1. We used the raw SDSS galaxies data described earlier in chapter 1. Galaxies are identified using the SDSS data processing pipeline. The data files contain the right ascension, declination and red shift along with other data for use in different kinds of research. For our purposes, these three fields are the important ones. These data are a representation of the distribution of galaxies on the celestial sphere. We then pixelize this data into matrices that span right ascension and declination in a specific redshift range.
We use these matrices as inputs to the signal matrix to perform the wavelet packet transform. The next step is we estimate the power spectrum, we radially average the power spectrum for convenient purposes. If fractals are present, the power spectrum should have a power law behaviour. To extract the power law we fit a line to log-log plot of the power spectrum vs the frequency curve, and calculate the slope of that line. Depending on the value of the slope, we estimate the value of the Hurst exponent $H$. Finally, we calculate the fractal dimension $D$ from the Hurst exponent. This is a general description of the workflow. In the following sections we more fully explain our procedures.

![Work flow](image)

Figure 3.1: Workflow of the entire process.

### 3.2 Reproducing the Distribution of Galaxies

We cannot project the galaxy locations from the celestial sphere directly onto a two dimensional matrix because that would result in distortions. To place the galaxy distribution into a form that can be accurately represented in a matrix, we used the HealPix technique, which we will describe in the following sections.
3.2.1 Raw SDSS Data

Here we discuss specific details of how we handled the SDSS data for the particular research we performed. The BOSS survey consists of the locations of galaxies in the celestial sphere as given by the coordinates of right ascension (RA), declination (DEC), redshift (z), and weights. The weight is a measure of how well identified the object is; a weight of 1 means that the SDSS pipeline determined the object is a galaxy with 100% certainty. The redshifts were eventually converted to comoving distances. The real survey consists of around 600,000 galaxies. The Mock Catalogs Galaxies had the same basic parameters but the number of galaxies were significantly larger; around 33 million galaxies [25].

3.2.2 Creating the Matrices using HealPix and SDSSPix

By the 90s as the cosmic microwave background measurements were a reality, researchers faced enormous challenges in analyzing large data sets on a celestial sphere [17]. To handle this, the HealPix algorithm was developed. HealPix (the Hierarchical Equal Area iso-Latitude Pixelization) was first developed for use with NASA’s COBE mission. HealPix is now a standard method of pixelizing data obtained from the celestial sphere. According to the NASA Jet Propulsion Laboratory, the major requirements in the development of HealPix were: to develop a mathematical framework which supports a suitable discretization of functions on a sphere at sufficiently high resolution, and to facilitate fast and accurate statistical and astrophysical analysis of massive full-sky data sets. The nature of the SDSS survey and the underlying spherical geometry of the survey allowed us to implement the HealPix algorithm in creating the 2-d matrices. SDSS designed a more specialized version of the HealPix to match the survey coverage called the SDSSPix. Essentially, the two schemes are the same except that the SDSSPix employs offsets to account for the survey geometry SDSS uses.

In both versions, the sky is divided into equal-area pixels. This process of pixeliza-
tion produces a subdivision of a spherical surface in which each pixel covers the same surface area as every other pixel. Figure 3.2 below shows the partitioning of a sphere at progressively higher resolutions. As we can see from the figure, the higher the resolution, the higher the number of pixels and the smaller is the size of the pixel. Areas of all pixels at a given resolution are identical [16]. These pixels are indexed according to the scheme illustrated by figure 3.3. There are two numbering schemes known as: the ring scheme, and nested scheme. In both situations, the two dimensional distribution of discrete area elements are mapped on a sphere into the one dimensional, integer pixel number array. The interested reader is referred to [17] for more details about this indexing process. Following this process of partitioning the sky into equal pixels of equal areas, and indexing all these pixels, we can project our data from the spherical surface onto a 2-d matrix.

Figure 3.2: The celestial sphere is divided into pixels of equal sizes. Depending on the desired resolution, the partitioning of the sphere is varied into low resolution or high resolution grids. The sphere on the upper left has the lowest possible resolution with 12 pixels of equal sizes and equal areas. As we move clockwise, the resolution is higher and the number of pixels is larger. The resolution of the sphere on the lower left is the highest with 768 pixels. Source: reproduced from the NASA Jet Propulsion Laboratory. https://healpix.jpl.nasa.gov
Figure 3.3: The two natural pixel indexing schemes, ring and nested. From top to bottom: Panel one (resolution = 2) and panel two (resolution = 4) show the ring scheme. Panel three (resolution = 2) and panel four (resolution = 4) show the nested scheme. Source: reproduced from the NASA Jet Propulsion Laboratory. https://healpix.jpl.nasa.gov

To express this process mathematically, we will introduce the special set of coordinates the SDSS developed. The new coordinates are called: LAMBDA (\( \lambda \)) and ETA (\( \eta \)), and they are similar to the celestial coordinates: right ascension (RA) and
declination (DEC). To convert between the RA, and DEC coordinates to the $(\lambda)$, and $(\eta)$ coordinates [18]

$$x = \cos(RA - node) \cos(DEC)$$ (3.1)

$$y = \sin(RA - node) \cos(DEC)$$ (3.2)

$$z = \sin(DEC)$$ (3.3)

And for $\lambda$, and $\eta$

$$\lambda = - \arcsin(x)$$ (3.4)

$$\eta = \arctan \left( \frac{z}{y} \right) - \eta_{pole}$$ (3.5)

According to the SDSS, the constants $node = 32.5$, $\eta_{pole} = 95.0$ are fixed throughout the survey.

To convert from the celestial coordinates $\lambda$, and $\eta$ above to the pixel index in SDSSPix, we first use the following intermediate formulae

$$i = \left\lfloor \frac{n_x \times \eta}{2\pi} \right\rfloor$$ (3.6)

where $n_x$ is defined according to the SDSS as the resolution times the number of pixels in the $x$ direction while,

$$j = \left\lfloor n_y \times \frac{1 - \cos(\lambda)}{2} \right\rfloor$$ (3.7)

the $n_y$ is defined as the resolution times the number of pixels in the $y$ direction. Using $i$, $j$, and $n_x$, we obtain the HealPix pixel $h_{index}$, using

$$h_{index} = n_x \times j + i$$ (3.8)
Finally we convert the pixel index $h_{\text{index}}$ to matrix indices $(r, c)$ using the following equations

$$c = \left\lfloor \frac{h_{\text{index}}}{n_x} \right\rfloor + 1 \quad (3.9)$$

$$r = [h_{\text{index}}]_{n_z} + 5 \times res \quad (3.10)$$

where $r$, and $c$ corresponds to the location of the column and the rows in the matrix, respectively. The main MATLAB routine we used to do this pixelization is Datahandler. The inputs to that program are (RA, DEC, res). RA and DEC are the right ascension and declination coordinates of the corresponding galaxies extracted from the raw SDSS data at different redshifts. These redshifts were converted into proper distances. Extracting these coordinates and converting redshifts to proper distances was done by MATLAB routines. Finally, the input (res), is the desired resolution for analysis. As described, the higher the resolution, the finer the details, and the larger the size of the matrix. Out of the possible resolutions outlined in [18], we used resolutions (8, 12, 16, 32, 64). All the resolutions used cover the entire sky but as the resolution increases, the sky is divided into more pixels. As a result, the pixel size at high resolutions would be much smaller than the size at low resolutions. For example, the pixel size at resolution 64 is 0.003 deg$^2$ compared to a pixel size of 0.333 deg$^2$ at resolution 8. The main routine Datahandler consist of different subroutines as outlined below.

**Datahandler Routine** Datahandler main routine consists of the subroutines: (sdsspix), (radec to lambdaeta), and (matrix) subroutines. The primary goal of this process is to convert the raw (RA-DEC) coordinates provided in the SDSS datasets to SDSSPix indices for a given resolution. After that, it converts the indices into a two dimensional matrix. See (Appendices B, C, D) for details.
3.3 Calculating the Power Spectrum

After we create the 2-d matrices using the SDSSPix and HealPix algorithms, we use these matrices as an input in the Wavelet Packet Transform (WPT). The resulting wavelet packet coefficients are used to compute the power spectrum.

3.3.1 Estimating the Power Spectrum From Wavelets

There are many different methods to estimate the power spectrum of a density field. The Fourier transform is the most common method, but the use of wavelets to estimate the power spectrum has many advantages over the traditional methods as discussed in section 2.4. We will show how the machinery of estimating the power spectrum using the Fourier methods can be extended to the wavelets. In order to use any method to estimate the power spectrum, the method has to be of a lossless nature; that is, the power is preserved before and after the transformation. To prove this conservation of the power, the Parseval’s theorem has to be valid for that method. It is well known that the Fourier transform is unitary and preserves the power, that is “the sum (or integral) of the square of a function is equal to the sum (or integral) of the square of its transform.” For our purposes, we have to show that the Parseval’s theorem is valid for the new wavelet basis. Fang and Pando [19,20] proved the lossless nature of the Discrete Wavelet Transform DWT so it can be used to estimate the power spectrum. Consider a density contrast \( \epsilon(x) \) in the density field of some physical structure, i.e., galaxies, with respect to the mean density of the field \( \bar{\rho} \), that is given by

\[
\epsilon(x) = \frac{\rho(x) - \bar{\rho}}{\bar{\rho}} \tag{3.11}
\]

We can write the Fourier expansion of the density contrast \( \epsilon \) as

\[
\epsilon(x) = \sum_{n=-\infty}^{\infty} \epsilon_n e^{i2\pi n x/L} \tag{3.12}
\]
The field can be Fourier transformed giving the coefficients $\epsilon_n$

$$\epsilon_n = \int_0^L \epsilon(x)e^{-i2\pi nx/L}dx$$  \hspace{1cm} (3.13)$$

Because both the DWT and Fourier have complete basis sets, any function can be represented by either bases [19]. By using the Parseval’s theorem, we can establish the relation between the coefficients of the Fourier expansion and the power imprinted in the density field by using the orthonormal Fourier basis functions as

$$\sum_{n=-\infty}^{\infty} |\epsilon_n|^2 = \frac{1}{L} \int_0^L |\epsilon(x)|^2dx$$ \hspace{1cm} (3.14)$$

The above equation is important for many reasons. First, it shows that the power is preserved under the transformation. Second, it relates the coefficients of the Fourier transform to the power of individual components in the frequency domain. Third, it establishes that the power spectrum is given by the Fourier coefficients squared. So the power spectrum of the perturbations is given by

$$P(n) = |\epsilon_n|^2$$ \hspace{1cm} (3.15)$$

Now, we can write the Parseval’s theorem with respect to the DWT bases as

$$\frac{1}{L} \int_0^L |\epsilon(x)|^2dx = \sum_{j=0}^{2^j-1} \frac{1}{L} \sum_{l=0}^{2^j-1} |\epsilon_{j,l}|^2$$ \hspace{1cm} (3.16)$$

where $\epsilon_{j,l}$ are the wavelet coefficients. So the power spectrum using the wavelet basis is defined as

$$P_j = \frac{1}{L} \sum_{l=0}^{2^j-1} |\epsilon_{j,l}|^2$$ \hspace{1cm} (3.17)$$

(See [19, 20] for the full proof). The above equation is important for two reasons. First, it shows that the use of the DWT to estimate the power spectrum is legitimate because it is of a lossless nature. Second, it shows also that the power spectrum is given by the coefficients squared. In our case, we have to prove the lossless nature of the Wavelet Packet Transform. By extending the case of the Fourier transform and the DWT to the WPT, we can write the Parseval’s theorem for the WPT as

$$\sum_{j=0}^{2^N-1} s_j \phi_{x_j, x_{j+1}}(x) = \sum_{j=0}^{2^{N-1}-1} a_j \phi_{x_j, x_{j+1}} + \sum_{j=0}^{2^{N-1}-1} d_j \psi_{x_j, x_{j+1}}$$ \hspace{1cm} (3.18)$$
The left hand side of the above equation represent the power before the transformation and it is equivalent to the left side of the Parseval’s theorem equation. The two terms on the right hand side represent the power after the transformation and it is equivalent to the summation on the right hand side of Parseval’s theorem equation. The energy balance between the original and decomposed signals is a direct consequence of Parseval’s theorem. Another way to view this is that wavelet packet decomposition is an orthonormal transformation, so the norm is preserved [33, 35, 36]. Thus, we have established that Parseval’s theorem holds true for the WPT, so we can use it to estimate the power spectrum. The power spectrum using the WPT is given by [28]:

\[ P_i = \text{var}[C_i, k] \] (3.19)

where \( \text{var}[C_i] \) is given by

\[ \text{var}(C_i, k) = \frac{1}{2^j} \sum_{k=0}^{2^j-1} |C_i, k|^2 \] (3.20)

and \((C_i, k)\) are the wavelet packet coefficients at level \(i\) and frequency \(k\). While we have established the WPT coefficients give a power, we have not yet discussed the frequency at which that power occurs. We will discuss the frequency details in the following section.

3.3.2 Frequency of the Signal

Natural Ordering

As we discussed in chapter 2, the WPT is the action of two filters on the signal. The two filters are known as low pass filter \((H)\), and the high pass filter \((G)\). The two filters decomposes the signal into high and low frequency components. The low pass filter \(H\) is associated with the approximation coefficients \(a_j\), and the high pass filter \(G\) is associated with the detail coefficients \(d_j\). The successive passes of the filters on the transformed coefficients \(a_j\) and \(d_j\) produces what is known as the wavelet packet tree as seen in figure 3.4.
Figure 3.4: The natural ordering of the WPT tree leaves. The $H$ is the low pass filter and $G$ is the high pass filter. The down arrow followed by number 2 means that the number of coefficients is reduced by 2 after each pass. The $\chi_{i,j}$ notation represents the transformed coefficients obtained from the pass $i$, with a wavelet packet node number $j$ [21].

In this wavelet packet tree, we have different leaves after each pass of the WPT. As the number of passes increases as we go down in the tree, the number of coefficients for each leaf is reduced by 2. Each leaf is characterized by two indices $\chi_{i,j}$. The index $i$ represents the level of wavelet decomposition, and the index $j$ represents the wavelet packet node number, or leaf, in the wavelet packet tree. This node is basically the output point of each filter after the action of the WPT on the transformed coefficients. For a 1-D signal, the number of these nodes is given by $2^n$ where $n$ is an integer corresponding to the level of decomposition. For example, at $n = 5$ we would have 32 wavelet packet coefficients produced at the corresponding 32 packet nodes. To further clarify, using the example of the leaf $\chi_{2,2}$, this corresponds to the approximation and detail coefficients obtained after the second pass, and has a node number of 2. The node numbers starts from 0 not 1 and increases from the
left to the right. This ordering of the leaves from the left to the right is known as the natural ordering. However, for reasons related to signal analysis explained in the following discussion, we are interested in the frequency ordering of the leaves, not the natural ordering.

**Frequency Ordering**

As introduced in chapter 2, the sweeps repeated on the approximation and detail coefficients with the high and low pass filters separate the various frequency components of the signal. As this process is repeated over and over again, every leaf in the wavelet packet tree will cover a narrower frequency band. As a result, the repeated passes of the two filters will mix the different types of coefficients and switch the high and low pass components. For example, the original signal corresponding to a frequency band 1 to 2 will be mirrored to another frequency band after the first pass because the low and high pass filters will act on both bands. This will mix the leaves in terms of what frequency band is really being spanned. Hence, it is important to associate the wavelet packet nodes to the frequency components they actually represent at each sweep. From the signal analysis perspective, the correspondence between the wavelet packet node number, the frequency order number (or the scale), and the spanned frequency band must be established.

To establish the relation between the natural and frequency ordering, we use the Gray Code numbering algorithm [21]. In the frequency ordering, the leaves are ordered in terms of the particular frequency band they represent, i.e., the increasing order of the frequency components from low to high, as corresponds to the low and high pass filters. In the Gray Code, it is a binary code in which two successive values differ only in one bit. Therefore, we can use it to re-order the packet nodes according to the frequency bands they represent. The following example illustrates how this reformulation works. Consider the number \( n = 5 \), the binary representation of it is given by [21]

\[
5 = (b_1, b_2, b_3, b_4) = 0111
\]
The Gray Code permutation for a number \( n \) is defined by the following equation

\[
GC(b_i) = (b_i + b_{i+1}) \mod 2
\]  
\[(3.22)\]

The natural order of an 8 wavelet packet nodes is given by

\[
0 1 2 3 4 5 6 7
\]  
\[(3.23)\]

The binary representation of that order is

\[
0000 \ 0001 \ 0010 \ 0011 \ 0100 \ 0101 \ 0110 \ 0111
\]  
\[(3.24)\]

The above representation can be written using the Gray Code permutation as

\[
0000 \ 0001 \ 0011 \ 0010 \ 0110 \ 0111 \ 0101 \ 0100
\]  
\[(3.25)\]

Converting the Gray Code above into decimal numbers, we get the wavelet packet node frequency ordering

\[
0 \ 1 \ 3 \ 2 \ 6 \ 7 \ 5 \ 4
\]  
\[(3.26)\]

Comparing the equation above with equation 3.23, we see the difference between the two numbering orders. In the first equation, the nodes are numbered in increasing order from the left to the right. However, the nodes are numbered in increasing order of the frequency content. In figure 3.5, the relation between the natural ordering and the frequency ordering is illustrated.
Figure 3.5: The WPT frequency tree for level of decomposition $n = 4$. Here $H(z)$ and $G(z)$ denote the low and high pass filters, respectively. The down arrows means that the number of coefficients is reduced by 2. The coefficients of the wavelet packet transform are ordered by the increasing order of frequency using the Gray Code permutation [21].

The range of frequencies is limited by the Nyquist frequency $0.5f_s$, where $f_s$ is the sampling frequency, as it is the case with other methods of spectrum estimation.

**The Spectrum in One and Two Dimensional Data Sets**

So by following all the steps outlined in previous sections, we can finally estimate the power spectrum in the case of the one dimensional WPT as

$$P_k = \text{var}(C_i, k)$$  

(3.27)

where $P_k$ is the power spectrum at frequency band $k$, and $(C_i, k)$ are the frequency ordered wavelet packet coefficients at the leaf $k$ of the wavelet packet tree. For the two dimensional WPT, we extend the techniques applied to the one dimensional...
WPT with the following modifications. In comparing the one dimensional WPT to the two dimensional case, the number of the wavelet packet leaves is increased by a factor of 4 in case of the 2-d WPT instead of 2 for the 1-d WPT. However, the number of coefficients for each leaf is reduced by a factor of 4 for the 2-d WPT instead of 2 as it is the case with the 1-d WPT. The reasons for this are related to the two nature of the WPT. First, in the 2-d WPT, the signal is represented by an \( m \times n \) matrix instead of a vector, so the transform is applied on both the rows and columns of the matrix. Secondly, in the case of 2-d WPT, there are two more types of coefficients (the horizontal difference \( h_i \) and the vertical difference \( v_i \) coefficients) in addition to the usual approximation and detail coefficients \( (a_i, d_i) \), and each one of them is represented by a packet leaf. By using the same natural ordering scheme, the leaves are numbered in increasing order from the left to the right starting from 0 to \( 2^j \) where \( j \) is the pass or the sweep number. In the case of the two dimensional WPT, the signal is represented by a \( 2^n \times 2^n \) matrix, instead of the signal vector in the case of the 1-d WPT. For the frequency ordering of the wavelet packet nodes, we apply the same Gray Code numbering scheme. The difference, however, is that we extend the sampling frequencies to be \( F_x \) and \( F_y \). To do this, we create a 2-d frequency grid that extends between 0 to 0.5\( f_x \) along the x-direction, and 0 to 0.5\( f_y \) along the y-direction. We can estimate the two dimensional power spectrum as

\[
P_{k_x,k_y} = \text{var}(C_{i(x,y)})
\]  

where \( P_{k_x,k_y} \) is the two dimensional power spectrum with respect to the frequencies \( k_x \) and \( k_y \), \( (C_{i(x,y)}) \) are the wavelet packet coefficients of the leaf \( i \) with indices \( x \) and \( y \), i.e, row and column locations in the 2-d matrix.

3.3.3 Radial Averaging

It is much more convenient to express the Cartesian 2-D power spectrum as a radial power spectrum. That is, we calculate the radial average frequency along the \( x \) and
Doing the same with the power spectrum, we average $P_k$ at the $i_{th}$ leaf in each of the two dimensional frequency grids

$$P_{avg}(k_i) = \sqrt{P_x^2 + P_y^2} \quad (3.30)$$

Powerspectrum is the main MATLAB routine that performs the wavelet packet decomposition outlined in chapter 2. The main inputs to the routine are (matrix, level, wavelet). The input, (matrix), is the two dimensional signal matrix from the first step of our analysis. The input, (level), corresponds to the wavelet decomposition level, or sweep number, $j$ as outlined in chapter 2. The choice of this number is tricky, and one has to consider different factors in picking a suitable $j$. For our purposes, we examined different values of $j$ such as (3, 4, 5, 6). Recall from chapter 2 that, as the number $j$ increases, the number of coefficients is lowered by a factor of 2, in comparison to the number of coefficients obtained at the previous $j$. As the number of coefficients decreases, small $n$ effects become more pronounced and the variance of a small number of coefficients becomes unstable. For that reason we have not considered any $j$ greater than 6. In deciding which $j$ to proceed with for the rest of the analysis, we examined the resulted values of the $R^2$ parameter (goodness of fit) produced from fitting a line to the log-log plot of $P(k)$ vs $k$. The slopes were consistent regardless of $j$, however, the goodness of fit varied somewhat. We had to choose a sweep that gave enough frequency detail, but did not suffer from low $n$ effects. The sweep number $j = 4$ gave us good $R^2$ and frequency resolution, so we decided to keep it fixed throughout the rest of the analysis. Finally, the input (wavelet) is the particular wavelet type we want to choose. Out of the possible options that MATLAB offers, we decided to choose the Daubechies1 wavelet, $db1$. The wavelet choice had almost no effect except at the boundaries. Since the $db1$ wavelet is the most compact, these effects were minimized using this wavelet.
Powerspectrum Routine performs the wavelet packet decomposition using the details described. In this routine, several MATLAB built-in functions were used such as (wpdec2, wpcoef). The (wpdec2) function performs the wavelet packet decomposition at two dimensions as its name implies. The (wpcoef) function give the option to extract information from the wavelet packet tree. The frequency ordering of the wavelet packet nodes using the Gray Code scheme outlined was done also within the main routine. See (Appendix A) for details. Finally, the output of the Powerspectrum routine is a plot of the radially averaged one dimensional power spectrum against the frequency, $P(k)$ vs. $k$. This plot is used as an input in the next step to estimate the Hurst exponent.

3.4 Estimating the Hurst Exponent $H$

3.4.1 On the Meaning of the Hurst Exponent

The Hurst Exponent $H$, also known as the index of dependence or the “index of long-range dependence” is a dimensionless parameter that indicates self-similarity present in the auto-correlation of a data set. Thus, it can also be defined as a measure of the self-similarity of a time series, or the clumpiness in a data set. It was first developed by Harold Edwin Hurst (1880-1978) to predict the flooding rate in the Nile River prior to the construction of the Aswan High Dam in Egypt. To understand this parameter better, we can look back at the original work of Hurst. The levels of rainfall in Central Africa appeared to be random each year, however, the flows of the Nile River seemed to show autocorrelation. That is, there is an inherited dependence between the rain fall at the present moment and its level in the past. Hurst concluded that the rainfall in one time period is likely influencing the rainfall in subsequent periods of time. So, Hurst documented this long-range dependence and proposed a quantitative description for the phenomena. According to the original proposition, meaningful values of $H$ lie in the range $[0, 1]$. Hurst in his original paper [24] published in 1951 argued that
1. If $H = 0.5$, the series represents a self-determining process. That is, the present value of the series is not dependent on the past values of the series.

2. When the value of $H$ is between $0 < H < 0.5$, the series is anti-persistent. High values in the series will likely revert their values towards the mean value, and then reduce to lower values than the mean value. The same applies for low values in the series, they are likely to revert towards the mean and then will have higher values than the mean.

3. When the value of $H$ varies between $0.5 < H < 1$, the values of the series rise in an upward direction and fall towards a downward direction with respect to the mean value.

Hurst developed the Rescaled Range Analysis method to calculate $H$. To learn more about the historical development of the Hurst exponent, see [39]. Long term memory research saw revolutionary developments during the the 70s and 80s. Now, there are different methods for estimating the Hurst exponent, and it has different applications in fields like medicine, finance, geology, and others. In our particular case, we used the wavelet packet method to estimate the Hurst exponent as discussed in the next section.

3.4.2 Estimating the Value of the Hurst Exponent

**Fitting a Line to the Log-Log Plot**

To find the Hurst exponent from the wavelet packet transform, we follow the procedures outlined in [23], which can be summarized by the following steps:

1. Estimate the power spectrum of the two dimensional data set, and plot the power spectrum $P(k)$ vs. the frequency $k$.

2. If the relation between the power spectrum and the frequency obeys the power law distribution, plot a modified log-log plot of $P(k)$ vs. $k$. 
3. Fit a line to the log-log plot of $P(k)$ vs. $k$.

4. Calculate the value of the slope of that line $\alpha$.

5. Estimate $H$ from the slope $\alpha$.

Figure 3.6 illustrates the fitting of a line to the log-log plot of $P(k)$ vs. $k$. The left panel shows the power spectrum vs frequency plot. The figure strongly hints at a power law dependence, possibility indicating the existence of fractals. The right panel is a log-log plot of the power spectrum. While noisy, for at least part of the frequency range, the data does appear generally linear. We used the correlation coefficient $R^2$ as a measure of the goodness of the linear fit to the log-log plot. Figure 3.7 shows the plot of the power spectrum against the physical scale and it also hints at a power law dependence.

Figure 3.6: On the left, the plot of the power spectrum vs. the frequency obeys the power law distribution. To the right, the log-log plot of the power spectrum vs. the frequency, with line fitted to the points.
Hurst Exponent from the Slopes

After fitting the log-log plot with a line, we find the slope of that line, \( \alpha \). Depending on the value of the slope, we use different formulae to calculate \( H \) from \( \alpha \). According to [23], if the slope of the line is between \( \alpha = (-1, -3) \), so \( H \) is given by

\[
H = \frac{\alpha + 3}{2} \tag{3.31}
\]

If the slope is bound between \( \alpha = (-1, 1) \), \( H \) is given by

\[
H = \frac{\alpha + 1}{2} \tag{3.32}
\]

If the slope is bound between \( \alpha = (1, 3) \), \( H \) is given by

\[
H = \frac{\alpha - 1}{2} \tag{3.33}
\]

With the Hurst exponent in hand, we are ready to finally calculate the fractal dimension \( D \). We will explore that in the next section.
## 3.5 Calculating The Fractal Dimension

The fractal dimension $D$ is related to the Hurst exponent $H$ by the relation

$$ D = n + 1 - H $$

(3.34)

where $n$ is known as the embedding dimension, i.e., (the minimum dimension where the object can lives in). Since we are treating galaxies as points (dimension 0) our embedding dimension is $n = 1$. In our analysis, we will slice the galaxy distribution into redshift (or comoving) spherical shells and compute the angular fractal dimension as a function of cosmological comoving space. The term angular fractal dimension is used because the analysis is being done on spherical shells with no depth. Thus, the roughness of the distribution as measured by the Hurst exponent and fractal dimension represent roughness across the spherical shell.

To recap, the procedure for calculating the fractal dimension is:

1. Using the HealPix and SDSSPix schemes, create the 2-d matrices that represent the galaxies distribution on the celestial sphere.

2. Apply the WPT on $2^n \times 2^n$ squared sampled matrices, and estimate the radially averaged power spectrum

$$ P_i = \text{var}[C_i] \Rightarrow \text{var}(C_i, k) = \frac{1}{2^j} \sum_{k=0}^{2^j-1} |C_i, k|^2 $$

(3.35)

3. Plot the log-log plot of the power spectrum vs. the frequency, and fit a line to the curve over the appropriate frequencies,

4. Calculate the Hurst exponent $H$ from the slope of the line $\alpha$. If $\alpha = (-1, -3)$, so $H$ is given by

$$ H = \frac{\alpha + 3}{2} $$

(3.36)

If $\alpha = (-1, 1)$, so $H$ is given by

$$ H = \frac{\alpha + 1}{2} $$

(3.37)
5. From the Hurst exponent $H$, calculate the fractal dimension by

\[ D = n + 1 - H \]  \hspace{1cm} (3.38)

In our case, equation 3.38 takes the form, $D = 2 - H$.

**Results Routine**  The Results routine calculates the Hurst exponent based on the values of the slope of straight line. The main inputs for this routine are (slopes 1, err 1, slopes 2, err 2). The inputs (slopes1, slopes2) are the slopes obtained from the log-log plots of $P(k)$ vs. $k$ for the BOSS survey galaxies, and the SDSS Mock Galaxy Catalogs, respectively. The other two inputs (err 1, err 2) are the confidence intervals of slopes1 and slopes2, respectively. The output of this routine are the corresponding Hurst exponents for both the BOSS galaxies and the Mock Galaxy Catalogs. They are represented as the plot of the proper distance vs. Hurst exponent as we will see in the next chapter.

The final step in our analysis is calculating the fractal dimension from the Hurst exponent. The outputs are the $(D_1, D_2)$ which corresponds to the fractal dimension of the BOSS galaxies and the Mock Galaxy Catalogs, respectively. See (Appendix F) for details. In the next chapter, we will present our results.
CHAPTER 4

Results and Interpretations

“...I do not think explicable by mere natural causes, but am forced to ascribe it to the counsel and contrivance of a voluntary agent.”

Isaac Newton’s letter to Richard Bentley, Dec 10, 1692

In this chapter, we will present the results of our analysis. In section 4.1, we discuss different procedures we followed and present our results. In section 4.2, we will interpret the results.

4.1 Results

In this section, we present the results of the analysis of the data using the entire outlined procedures. In particular, we applied the analysis on both the BOSS Galaxies and the Mock Galaxy Catalogs of the SDSS. We used five different resolutions in creating the two dimensional matrices and one particular type of wavelet, $db1$. At all resolutions, we sampled the original $m \times n$ matrix into $2^n$ square matrix at different locations on the celestial sphere, calculated the power spectrum and the slope at those locations, and averaged the results.

In dealing with the issue of edge distortions, MATLAB offers different ways to extend the signal at the boundaries. After trying out the different options, we settled down to the simplest choice: symmetric extension. In this scheme, the signal at these boundaries is replicated by symmetric values from the original signal. In other words, the signal at these boundaries is created by using constant values of the original signal. Although this extension scheme has one major drawback of producing artificial discontinuities, this effect was neutralized when we reduced
the original matrix into an $2^n$ square matrix. Besides that, the powerful localization property of wavelets kept the error in the signal, if any, extended to a few coefficients. It should be noted that any extension scheme will have its drawbacks.

In our analysis, we used the proper distance instead of the redshift to handle the data properly. We computed the proper distance at corresponding redshifts. In reproducing the distributions of galaxies, we used a redshift bin width of $z = 0.04$ which corresponds to (167.9885) Mpc after testing different bin widths such as $z = 0.01$ and $z = 0.1$. This bin width was optimal for our analysis because we treat galaxies as points and it takes into consideration the isotropy and homogeneity of the Universe. The bin width $z = 0.1$ which corresponds to (414.6571) Mpc takes into consideration the isotropy and homogeneity of the Universe; but it treats galaxies as three dimensional distribution, so it was not a good choice for our analysis. At the other bin width $z = 0.01$ which corresponds to (42.2576) Mpc, we had very few galaxies; and the isotropy and homogeneity of the Universe does not hold. At low redshift like $z = 0.4$ (corresponds to proper distance of $1.5473 \times 10^3$ Mpc), this bin width would correspond to a particular distance in Mpc (167.9885) Mpc. However, because the redshift-distance relation is not linear, so as $z$ increases slightly, the corresponding proper distance increases significantly. At higher redshifts like $z = 0.75$, using the same redshift bin width of 0.04 would correspond to a much higher distance in Mpc. Therefore, by using the proper distance instead of redshift, we ensure that the bin width is the same regardless of the redshift used. Hence, the use of proper distance ensured that our galaxy distribution is a realistic representation of the celestial sphere. In computing the proper distance from the redshift, we used the formula [1]

$$\text{d}_p(t) = a(t) \int_0^r \text{dr}$$

(4.1)

where $\text{d}_p$ is the proper distance, $a(t)$ is the scale factor, $\text{dr}$ is the radial comoving coordinate. The lower and upper limits of the integration are $\frac{1}{(1+z)}$, and 1 respectively.

In our analysis, we used the $\Lambda$ CDM model to be consistent with the SDSS Mock Galaxies Catalogs. We used the following parameters: $\Omega_m = 0.27$, $H_0 = 70.8$ km
$s^{-1}$ Mpc$^{-1}$, and $\Omega_\Lambda = 0.73$. To do this conversion, we used the routine (PropDist). See (Appendix E) for details.

In fitting the line, we used the correlation coefficient $R^2$ to determine the goodness of the fit. Any line with $R^2 < 0.8$ was not used in any further calculations. At all resolutions, the first and last point in the plot of power spectrum vs. frequency corresponding to $z = 0.4$ and $z = 0.75$ respectively had $R^2 < 0.8$; hence, these points were excluded from further consideration. Deeper analysis of the data at these redshifts is required to determine if the data have fractal behaviour. However, the lack of galaxies at these redshifts make this analysis currently unfeasible. The other redshift ranges had $R^2 > 0.8$. Finally, after calculating the different values of $H$ at different resolutions, we calculated the fractal dimension using the relation $D = 2 - H$.

We present our results as tables of slopes, figures of the Hurst exponent and fractal dimension plotted against the proper distance. The slopes at resolution 8 are shown in table 4.1

<table>
<thead>
<tr>
<th>BOSS Galaxies Slopes</th>
<th>Mock Catalogs Galaxies Slopes</th>
<th>Redshift Z</th>
<th>Proper Distance (Mpc $10^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.4422</td>
<td>-2.0895</td>
<td>0.45</td>
<td>1.7196</td>
</tr>
<tr>
<td>-1.6316</td>
<td>-2.1514</td>
<td>0.5</td>
<td>1.8871</td>
</tr>
<tr>
<td>-1.7904</td>
<td>-2.2163</td>
<td>0.55</td>
<td>2.0501</td>
</tr>
<tr>
<td>-1.8889</td>
<td>-2.1202</td>
<td>0.6</td>
<td>2.2086</td>
</tr>
<tr>
<td>-1.7056</td>
<td>-2.1441</td>
<td>0.65</td>
<td>2.3626</td>
</tr>
<tr>
<td>-1.4</td>
<td>-2.0825</td>
<td>0.7</td>
<td>2.5124</td>
</tr>
</tbody>
</table>

Table 4.1: Slopes of BOSS and Mock Catalogs Galaxies at resolution 8

and the Hurst exponent and fractal dimension are shown in figures 4.1 and 4.2 respectively
The Hurst exponent for the BOSS Galaxies starts with $H = 0.7$ and drops as the proper distance increases, then, it increases again to $H = 0.8$ at the highest proper distance. On average, the values lies with in the range $H = 0.6 \pm 0.2$. For the Mock
Catalogs Galaxies, the Hurst exponent is almost constant around $H = 0.4 \pm 0.1$. As seen from figure 4.1, $H$ for BOSS Galaxies is greater than Mock Catalogs Galaxies and there is no overlapping between the values. The fractal dimension for BOSS galaxies is $D = 1.4 \pm 0.2$ and it has the lowest value towards the highest proper distance. For Mock Catalogs Galaxies, the fractal dimension is almost constant around $D = 1.6$. There is no overlapping between the values of $D$. Based on the values of $H$ and $D$, the galaxy distribution at this resolution is fractal, irregular, and inhomogenous.

The slopes at resolution 12 are shown in table 4.2

<table>
<thead>
<tr>
<th>BOSS Galaxies Slopes</th>
<th>Mock Catalogs Galaxies Slopes</th>
<th>Redshift $Z$</th>
<th>Proper Distance (Mpc $10^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.04245</td>
<td>-1.6926</td>
<td>0.45</td>
<td>1.7196</td>
</tr>
<tr>
<td>-1.2764</td>
<td>-1.8359</td>
<td>0.5</td>
<td>1.8871</td>
</tr>
<tr>
<td>-1.3753</td>
<td>-1.8051</td>
<td>0.55</td>
<td>2.0501</td>
</tr>
<tr>
<td>-1.1964</td>
<td>-1.7072</td>
<td>0.6</td>
<td>2.2086</td>
</tr>
<tr>
<td>-1.0688</td>
<td>-1.7032</td>
<td>0.65</td>
<td>2.3626</td>
</tr>
<tr>
<td>-0.9631</td>
<td>-1.6643</td>
<td>0.7</td>
<td>2.5124</td>
</tr>
</tbody>
</table>

Table 4.2: Slopes of BOSS and Mock Catalogs Galaxies at resolution 12

and the Hurst exponent and fractal dimension are shown in figures 4.3 and 4.4 respectively
As the resolution is increased to 12, the Hurst exponent for BOSS Galaxies is about $H = 0.9 \pm 0.1$ and for Mock Catalogs Galaxies it is almost constant around $H = 0.6 \pm 0.1$. The Hurst exponent for BOSS Galaxies starts with high value as $H = 0.9$ and then decreases as the proper distance increases, then, it increases again toward
the highest proper distance. This scaling behavior of $H$ at resolution 12 is similar to that at resolution 8. For the fractal dimension, the Mock Catalogs Galaxies has a higher fractal dimension and it is almost constant around $D = 1.4 \pm 0.1$. For BOSS Galaxies, the fractal dimension starts as low as $D = 1.1$ and increases as the proper distance increases, then, it decreases again at the highest proper distance. On average, $D = 1.2 \pm 0.1$ for BOSS Galaxies. As it was the case with resolution 8, there is no overlap between the values of $H$ and $D$ in either type of galaxies. The values of $H$ and $D$ at this resolution indicates again that galaxy distribution is fractal, irregular, and inhomogenous.

The slopes at resolution 16 are shown in Table 4.3

<table>
<thead>
<tr>
<th>BOSS Galaxies Slopes</th>
<th>Mock Catalogs Galaxies Slopes</th>
<th>Redshift Z</th>
<th>Proper Distance (Mpc $10^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.7041</td>
<td>-2.0514</td>
<td>0.45</td>
<td>1.7196</td>
</tr>
<tr>
<td>-1.6704</td>
<td>-2.2551</td>
<td>0.5</td>
<td>1.8871</td>
</tr>
<tr>
<td>-1.7655</td>
<td>-2.2268</td>
<td>0.55</td>
<td>2.0501</td>
</tr>
<tr>
<td>-1.7556</td>
<td>-2.2339</td>
<td>0.6</td>
<td>2.2086</td>
</tr>
<tr>
<td>-1.8111</td>
<td>-2.1465</td>
<td>0.65</td>
<td>2.3626</td>
</tr>
<tr>
<td>-1.2879</td>
<td>-2.0539</td>
<td>0.7</td>
<td>2.5124</td>
</tr>
</tbody>
</table>

Table 4.3: Slopes of BOSS and Mock Catalogs Galaxies at resolution 16

and the Hurst exponent and fractal dimension are shown in figures 4.5 and 4.6 respectively.
As the resolution increases to 16, the Hurst exponent for BOSS Galaxies starts as high as $H = 0.7$ and remains relatively constant as the proper distance increases. On average, it is about $H = 0.7 \pm 0.1$. For Mock Catalogs Galaxies, it is almost constant around $H = 0.4$ except the first and last values which have slightly higher
$H$. On average, it is about $H = 0.4 \pm 0.1$. For the fractal dimension, the Mock Catalogs Galaxies has a higher fractal dimension and it is almost constant around $D = 1.6 \pm 0.1$. For BOSS Galaxies, the fractal dimension is almost constant about $D = 1.4 \pm 0.1$ except it decreased slightly at the highest proper distance to $D = 1.2$.

As it was the case with resolutions 8 and 16, there is no overlap between the values of either $H$ and $D$ in either type of galaxies. The values of $H$ and $D$ at this resolution indicates again that galaxy distribution is fractal, irregular, and inhomogenous.

The slopes at resolution 32 are shown in table 4.4

<table>
<thead>
<tr>
<th>BOSS Galaxies Slopes</th>
<th>Mock Catalogs Galaxies Slopes</th>
<th>Redshift $Z$</th>
<th>Proper Distance (Mpc $10^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.8091</td>
<td>-1.9772</td>
<td>0.45</td>
<td>1.7196</td>
</tr>
<tr>
<td>-1.3105</td>
<td>-2.4298</td>
<td>0.5</td>
<td>1.8871</td>
</tr>
<tr>
<td>-1.3138</td>
<td>-2.4522</td>
<td>0.55</td>
<td>2.0501</td>
</tr>
<tr>
<td>-1.2155</td>
<td>-2.4136</td>
<td>0.6</td>
<td>2.2086</td>
</tr>
<tr>
<td>-1.106</td>
<td>-2.2547</td>
<td>0.65</td>
<td>2.3626</td>
</tr>
<tr>
<td>-0.8871</td>
<td>-2.0039</td>
<td>0.7</td>
<td>2.5124</td>
</tr>
</tbody>
</table>

Table 4.4: Slopes of BOSS and Mock Catalogs Galaxies at resolution 32

and the Hurst exponent and fractal dimension are shown in figures 4.7 and 4.8

Figure 4.7: Hurst exponent at resolution 32
At resolution 32, the Hurst exponent exhibits a scaling behavior similar to the previous resolutions. For BOSS Galaxies, it starts as high as $H = 0.9$, then it decreases and remains relatively constant as the proper distance increases. Towards the highest proper distance, the value of $H$ increases again to $H = 0.9$. On average, $H = 0.8 \pm 0.1$. For Mock Catalogs Galaxies, the Hurst exponent starts as high as $H = 0.5$ and then decreases and remains relatively constant around $H = 0.3$ as the proper distance increases. Towards the highest proper distance, the value of $H$ increases again. On average, $H = 0.3 \pm 0.2$. The fractal dimension for BOSS Galaxies is relatively constant around $D = 1.1 \pm 0.1$. For Mock Catalogs Galaxies, it has a higher fractal dimension around $D = 1.7 \pm 0.2$. As it is the case with the previous resolutions, there is no overlapping between the values of either $H$ or $D$ in either type of galaxies. Again as it is the case with the previous resolutions, the values of $H$ and $D$ indicates fractal, irregular, and inhomogenous galaxy distribution.
The slopes at resolution 64 are shown in table 4.5

<table>
<thead>
<tr>
<th>BOSS Galaxies Slopes</th>
<th>Mock Catalogs Galaxies Slopes</th>
<th>Redshift Z</th>
<th>Proper Distance (Mpc (10^3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.52846</td>
<td>-1.3352</td>
<td>0.45</td>
<td>1.7196</td>
</tr>
<tr>
<td>-0.91399</td>
<td>-1.8775</td>
<td>0.5</td>
<td>1.8871</td>
</tr>
<tr>
<td>-0.98708</td>
<td>-1.9407</td>
<td>0.55</td>
<td>2.0501</td>
</tr>
<tr>
<td>-0.93711</td>
<td>-1.8187</td>
<td>0.6</td>
<td>2.2086</td>
</tr>
<tr>
<td>-0.7051</td>
<td>-1.6445</td>
<td>0.65</td>
<td>2.3626</td>
</tr>
<tr>
<td>-0.5527</td>
<td>-1.3694</td>
<td>0.7</td>
<td>2.5124</td>
</tr>
</tbody>
</table>

Table 4.5: Slopes of BOSS and Mock Catalogs Galaxies at resolution 64

and the Hurst exponent and fractal dimension are shown in figures 4.9 and 4.10 respectively

Figure 4.9: Hurst exponent at resolution 64
At the highest resolution of 64, the Hurst exponent exhibits a scaling behavior similar to the previous resolutions but the values shows much more variations. For BOSS Galaxies, it starts as high as $H = 0.85$, then it decreases and remains relatively constant as the proper distance increases. Toward the highest proper distances, the value of $H$ increases again to $H = 0.85$. On average, $H = 0.6 \pm 0.2$. For Mock Catalogs Galaxies, the Hurst exponent starts as high as $H = 0.7$ and then decreases significantly to reach $H = 0.2$ and $H = 0.1$ as the proper distance increases. Towards the highest proper distance, the value of $H$ increases again. It is hard to find an average value for $H$ in that case but roughly it is, $H = 0.4 \pm 0.3$. The fractal dimension for BOSS Galaxies is relatively constant around $D = 1.5 \pm 0.3$. For Mock Catalogs Galaxies, it has a higher value and an interesting behavior of $D$. It starts as low as $D = 1.3$ at the lowest proper distance and overlaps with the value of $D$ of BOSS Galaxies. The value of $D$ increases to be very near but not equal to $D = 2$ as the proper distance increases. Towards the highest proper distances the value of $D$ decreases to reach $D = 1.3$ again. As it is the case with the Hurst exponent, it is hard to find an accurate average but roughly, $D = 1.5 \pm 0.3$ would be a good
estimation. This interesting behavior of $D$ at resolution 64 implies that galaxies distribution is still fractal but they are starting to get smoother whenever $D \approx 2$. Again as it is the case with the previous resolutions, the values of $H$ and $D$ indicates fractal, irregular, and inhomogenous galaxy distribution.

4.2 Interpretation

Based on the values of Hurst exponent at different resolutions, they are in the range $(0.2 - 0.7)$, hence, it appears that there are different degrees of roughness. All the results indicate that the galaxy distribution is irregular and inhomogenous. In all of the resolutions used, all the values of $D$ reported were lower than 2, and some even close $D = 1$. As explained in chapter 1, systems like galaxies are said to have irregular distribution if the measured fractal dimension $D < n$. In our case, a $D < 2$ will indicate a fractal distribution.

We make the following conclusions: the angular distribution of galaxies in the redshift range $0 < z < 0.8$ are fractal systems; they have irregular distribution, and they are inhomogenous. Our results are in excellent agreement with previous work as described in section 1.4.3. Previous work in the fractal analysis of galaxies showed that galaxies can be described as fractal systems in the redshift range $0.1 < z < 5$, they are irregular, and they are inhomogenous. Our results agrees with these findings.

There are other features that are worth mentioning. One similarity is the values of Hurst exponent was consistent for both BOSS and Mock Catalogs Galaxies. The Hurst exponent for BOSS galaxies was always higher than Mock Catalogs Galaxies. Consequently, the fractal dimension for the Mock Catalogs Galaxies was higher than BOSS Galaxies. By higher we do not mean to imply the Mock Catalogs was “more” fractal. Indeed, the Mock Catalogs Galaxies had fractal dimensions closer to 2, indicating that these were nearly planarly distributed.

Another feature is that the chosen resolution does not play a role in producing the results. If our results were dependent on the chosen resolution, it would seri-
ously raise questions on our method. For example, if the Hurst exponent for BOSS Galaxies was higher than Mock Catalogs Galaxies at resolution 16 but it was lower at resolution 32, that would be interpreted as an effect of the resolution - not as part of any underlying fractal behaviour.

Another consistency in all the cases is that the Hurst exponent and fractal dimension values lie within the same range. The Hurst exponents are in the range of $H = (0.5 \pm 0.3)$ for BOSS Galaxies and $H = (0.4 \pm 0.2)$ for Mock Catalogs Galaxies. For the fractal dimension, the values of $D$ are always in the range of $(1 - 2)$ for both the BOSS and Mock Catalogs Galaxies. This consistency in ranges at all resolutions is another indication of the robustness of the statistics we used.
CHAPTER 5

Conclusions and Future Research Goals

“Solomon said, My Lord, enable me to be grateful for Your favor which You have bestowed upon me and upon my parents and to do righteousness of which You approve. And admit me by Your mercy into the ranks of Your righteous servants.”

Quran 27:19

In this final chapter, we restate our conclusions and compare them with the published literature. In doing so, we will emphasize on the successes and concerns of our research. We finish the chapter by discussing potential future research. Particularly, we will build our future research goals based on the reported results.

5.1 Conclusion

We developed a novel way of using wavelet packets to calculate the angular fractal dimension of galaxy distributions as a function of cosmological redshift. We used data from the BOSS survey and Mock Galaxies Catalogs of the SDSS. The fractal dimension as applied to galaxies is a measure of the degree of homogeneity or regularity in the structure. Values of $D < 2$ indicate an irregular and inhomogenous distribution, and values of $D = 2$ indicate regular and homogenous distribution. Using the HEALPix and SDSSPix schemes, we reproduced the galaxy distributions on the celestial sphere to create the two dimensional signal matrix. This matrix was subjected to the WPT to estimate the power spectrum. We fitted a line to the log-log plot of the power spectrum vs. frequency. We estimated the Hurst exponent
from the slope of the line and calculated the fractal dimension from the Hurst exponent. The Hurst exponent and fractal dimension are related to each other by the relation $D = 2 - H$; the higher the Hurst exponent the lower is the fractal dimension; the lower the Hurst exponent the higher is the fractal dimension. We reported values of Hurst exponent $H = (0.5 \pm 0.3)$ for BOSS Galaxies, and $H = (0.4 \pm 0.2)$ for Mock Galaxies Catalogs. The fractal dimension reported for BOSS Galaxies was $D = (1.3 \pm 0.2)$ and for Mock Catalogs it was $D = (1.6 \pm 0.2)$. In both cases, the fractal dimension was always lower than 2. The Hurst exponent for BOSS Galaxies was higher than Mock Catalogs Galaxies, hence, the fractal dimension for Mock Catalogs Galaxies was higher than BOSS Galaxies. The results are consistent at all resolutions which indicates the success of our method and the robustness of the statistics we used.

5.2 Goals of Future Research

The future research is guided by a number of goals that can be developed into separate projects. Firstly, testing our algorithm with data at high redshift, i.e $z > 1$. The BOSS survey we used covers redshift range between $0.8 > z > 0$, so we did not have data at high redshifts to test our method. We are looking forward to data from the Dark Energy Survey and other deep redshift surveys. Given that the Universe is smoothing out as the redshift goes higher, we would expect different values of $D$ as the redshift increases. A potential research project might be to use varying $D$ to probe if the Universe has entered different dynamical epochs. Because fractal behavior provides us with an indication about linear and non-linear regimes, it may be possible to demarcate the point at which dark energy becomes the dominant component governing the evolution of the Universe.

Another possible goal is more inclined towards theoretical cosmology rather than observational cosmology. An interesting problem that has recently seen vigorous activity is the scale of homogeneity in the Universe. While it has been taken for
granted that the Universe is homogenous and isotropic at large scales, recent investigations have suggested that the *observational* homogeneity has to be defined before we make such a conclusion. This problem was approached recently [12] by the fractal analysis of large scale structure. The authors have concluded that the observational homogeneity is interconnected with the fractality of structure. What we are looking forward to is to test out this fundamental relation between the fractal nature and the homogeneity of the structure. Is there is any meaningful difference between the two? Does the fractality of the structure extend as far as the limit of the observational homogeneity, or the fractality breaks out at some scale? What is that scale and why this happens? We will look forward to combine our theoretical models with the analysis of observational data to find answers to all of these questions in the future.
Appendices
APPENDIX A

Program: Powerspectrum

function finpower=powerspectrum(data)

jn=4; % the depth (no. of sweeps=5)

p=0:1:2^jn-1;

[fv1,fv2] = meshgrid(p/2^jn);

n=bitxor(p,fix(p/2));

tab=[ 1 2; 3 4];

for u = 1:jn - 1

    tab= [tab tab+4^u];

end

for v=1:jn -1

    tab=[tab; tab+2*4^v];

end
T = wpdec2(data,jn,'db1');

PS=zeros(2^jn,2^jn);

for p1 = 1:2^jn
    for p2 = 1:2^jn
        Cjk=wpcoef(T,[jn -1+tab(n(p1)+1,n(p2)+1)]);
        PS(p1,p2)=var(Cjk(:)); % power spectrum
    end
end

[kx,ky]=size(PS);

[~,rho]=cart2pol(1:kx,1:ky);

minrho=1/rho(2);

rho=rho.*minrho;

sz=length(PS);

rho=round(rho);

i=cell(sz+1,1);
radavgps=zeros(sz+1,1);

eradavgps=zeros(sz+1,1);

for r=0:sz
    ir+1=find(rho==r);
end

for r=0:sz
    radavgps(r+1)= nanmean(PS(ir+1));
    eradavgps(r+1)= nanstd(PS(ir+1));
end

xfreq=diag(fv1);

yfreq=diag(fv2);

avgfreq=sqrt(xfreq.^2+yfreq.^2);

finfreq=avgfreq(1:16)

finpower=radavgps(1:16)

mdl=fitlm(log(finfreq),log(finpower))

finpower(isnan(finpower(:,1)),:)=[]
finpower = finpower(2:8)

newlen = length(finpower)

finfreq = finfreq(1:newlen)

plot(log(finfreq), log(finpower), '.k', 'MarkerSize', 18)

toc
end
function data = datahandler(ra,dec,res)
%datahandler takes ra and dec of galaxies and a given resolution, res and
% places them in a matrix in which each element is equal area.
% Based on HealPiX ideas
% Uses routines myhealpix and mymatrix
% Output: data is the matrix to be fractal analyzed.
pixels = myhealpix(ra,dec,res);
data = mymatrix(pixels,res);
end
APPENDIX C

Program: myhealpix

function [ index] = myhealpix(ra,dec,res)
%myhealpix takes ra, dec, res and returns a pixel number based on
% healpix ideas and SDSS geometry.

nx0 = 36;
ny0 = 13;

d2r = pi/180;
r2d = 180/pi;

etaOffSet = 91.25;
surveyCenterRA = 185.0;
surveyCenterDEC = 32.5;
node = d2r*(surveyCenterRA - 90.0);
etaPole = d2r*surveyCenterDEC;
% Converting RA-DEC -> LAMBDA-ETA
[lambda,eta]=radec_to_lambdaeta_SDSSPIX(ra,dec);

% Converting LAMBDA-ETA -> HEALPix index
nx = nx0*res;
ny = ny0*res;

pixnum = eta;

for n = 1:length(eta)
et1 = (eta(n) - etaOffSet) * d2r;

if et1 >= 0
    et2 = et1;
else
    et2 = et1 + 2*pi;
end

i = floor((nx*et2)/(2*pi));

lambda2 = (90 - lambda(n)) * d2r;

if lambda2 >= pi
    j = ny - 1;
else
    j = floor(ny*((1-cos(lambda2))/2));
end

pixnum(n) = nx*j+i;
end

index = pixnum;
end
function [ matrix ] = mymatrix(pixels, res)
% mymatrix takes pixels and puts them in a matrix

nx0 = 36;
y0 = 13;

nx = nx0 * res;
y = ny0 * res;

matrix = zeros(nx, ny);

for n = 1:length(pixels)
    column = floor(pixels(n)/nx)+1;
    row = mod(pixels(n), nx) + 5*res;
    if row > nx
        row = row - nx;
    end
    matrix(row, column) = matrix(row, column)+1;
    clear row;
    clear column;
end
end
APPENDIX E

Program: PropDist

function pd = PropDist(z)

% this function is used to determine the proper distance of a galaxy
% based on its red shift.
lowerlimit = 1./(1 + z);
upperlimit = 1;
integrand = @(a) 1./sqrt((0.27.*a) + (0.73.*(a.^4)));

c = 2.99792458E8;
H = 70.8E3;

pd = quadgk(integrand, lowerlimit, upperlimit).*((c./H));

end
function [hurst1,hurst2,fracdim1,fracdim2] = results(slopes1,fit1,slopes2,fit2)
properdistance=[1.7196;1.8871;2.0501;2.2086;2.3626;2.5124]*10^3;
    for i=1:length(slopes1)
        if (slopes1(i)>= -1)
            hurst1(i)=(slopes1(i) +1)/2;
        elseif (slopes1(i) < -1)
            hurst1(i)= (slopes1(i) +3)/2;
        end
    end
    for j=1:length(slopes2)
        if (slopes2(j)>= -1)
            hurst2(j)=(slopes2(j) +1)/2;
        elseif (slopes2(j) < -1)
            hurst2(j)=(slopes2(j) +3)/2;
        end
    end

errorbar(properdistance,hurst1,fit1,.k','MarkerSize',18); hold on
xlabel('Proper Distance','FontSize',14,'FontWeight','bold','Color','k')
ylabel('Hurst Exponent H','FontSize',14,'FontWeight','bold','Color','k')
xlim([1.5 2.8]*10^3)
ylim([0 1.1])

errorbar(properdistance,hurst2,fit2,.b','MarkerSize',18)
fracdim1 = 2 - hurst1
fracdim2 = 2 - hurst2

figure
errorbar(properdistance, fracdim1, fit1, '.k', 'MarkerSize', 18); hold on
errorbar(properdistance, fracdim2, fit2, '.b', 'MarkerSize', 18)
REFERENCES


galaxy distribution in the redshift range $0.45 < z < 5.0$. arXiv:1409.5409


[22] Zhiguo Wang ; Dechun Guo ; Xi Li ; Yuanchun Fei. Estimating Hurst exponent with wavelet packet. 2006 7th International Conference on Computer-Aided Industrial Design and Conceptual Design


[34] C. Sidney Burrus. Generalizations of the Basic Multiresolution Wavelet System. OpenStax-CNX module: m45095


[39] Graves et.al. A brief history of long memory: Hurst, Mandelbrot and the road to ARFIMA