Genomic Signal Processing and Regulatory Networks:
Representation, Dynamics and Control

BY

LIMING WANG
B.S., Huazhong University of Science and Technology, China, 2006

THESIS
Submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering in the Graduate College of the University of Illinois at Chicago, 2011

Chicago, Illinois

Defense Committee:

Dan Schonfeld, Chair and Advisor
Rashid Ansari
Daniela Tuninetti
Laura DeMarco, Mathematics, Statistics, and Computer Science
Shmuel Friedland, Mathematics, Statistics, and Computer Science
ACKNOWLEDGMENT

First and foremost, I would like to thank my advisor Prof. Dan Schonfeld. It is truly my fortune to have Dan as my advisor and friend, who has a broad view and insight of my research area, who is always supportive for my decisions and enthusiastic about my research work.

I also would like to thank all the members in my committee, Prof. Rashid Ansari, Prof. Laura DeMarco, Prof. Shmuel Friedland and Prof. Daniela Tuninetti for their time and suggestions.

Special thanks also go to Prof. Stefan Wenger who was my advisor in math department for all he taught me and his guidance.

I am grateful for all the faculty members in Department of Electrical and Computer Engineering and Department of Mathematics, Statistics, and Computer Science who taught me during my graduate studies. Thank you all for showing me the beauties of mathematics, which enlighten and serve as impetus for my research.

I wish to express my gratitude to all my friends who helped me and made my life in Chicago much more interesting, especially Yang Weng, Liuling Gong, Xiangqiong Shi, Xu Chen, Nicola Piotto, Lu Zheng, Junmei Wang, Nidhal Bouaynaya, Chong Chen, Hunsop Hong, Xiang Ma, Pan Pan, Junlan Yang, Songqing Zhao, Jing Huang, Qun Li, Yanying Chen, Stefano Rini and Ishita Basu.

As always, I would like to thank my parents for their unconditional love and care from afar.

LW
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>INTRODUCTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1 Overview of Genomic Signal Processing and Regulatory Networks</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.2 Genomic Signal Representation and Dynamics</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.2.1 Genomic Signal Representation</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.2.2 Dynamics in Genomic Signal Processing</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1.2.3 Control in Regulatory Network</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>1.3 Summary of the Contributions</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>MAPPING EQUIVALENCE FOR SYMBOLIC SEQUENCES</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>DYNAMICS, STABILITY AND CONSISTENCY FOR SYMBOLIC SEQUENCES PROCESSING</td>
<td>45</td>
</tr>
<tr>
<td>4</td>
<td>STOCHASTIC CONTROL FOR REGULATORY NETWORK</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>4.1 Introduction</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>4.2 The Models for Regulatory Network</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>4.3 The Single Player Case</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>4.4 The Non-cooperative Multiple Players Case</td>
<td>67</td>
</tr>
</tbody>
</table>
## TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>Solving Nash Equilibrium</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Approximate Nash Equilibrium</td>
</tr>
<tr>
<td>4.6</td>
<td>Experiments on Mammalian Network</td>
</tr>
<tr>
<td>4.7</td>
<td>Conclusion</td>
</tr>
<tr>
<td>5</td>
<td>CONCLUSIONS AND FUTURE WORK</td>
</tr>
<tr>
<td>5.1</td>
<td>Summary and Contributions</td>
</tr>
<tr>
<td>5.2</td>
<td>Recommendations for Future Work</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Consistency Criterion for Composition of Operators</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Correlated Equilibrium and New Solution Concept</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Switching Control Game and Super-modularity</td>
</tr>
</tbody>
</table>

## CITED LITERATURE

82

## VITA

88
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>74</td>
</tr>
<tr>
<td>I BOOLEAN FUNCTIONS OF MAMMALIAN CELL CYCLE</td>
<td>74</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>41</td>
</tr>
<tr>
<td>5</td>
<td>42</td>
</tr>
<tr>
<td>6</td>
<td>43</td>
</tr>
<tr>
<td>7</td>
<td>57</td>
</tr>
<tr>
<td>8</td>
<td>58</td>
</tr>
<tr>
<td>9</td>
<td>59</td>
</tr>
<tr>
<td>FIGURE</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>10</td>
<td>Inconsistency case: The illustration of how Euclidean distance for two mappings, for which one is in the previous Fatou component $U$ and the other is not, changes with the number of iterations for human gene AD169 sequence.</td>
</tr>
<tr>
<td>11</td>
<td>Illustrations of the average reward of the optimal intervention for control $a_1$ on gene Rb under the case only one control and the average reward of the optimal intervention for control on Rb using IMDP method with the other control $a_2$ on CycA ignored.</td>
</tr>
<tr>
<td>12</td>
<td>Illustrations of the average reward of the optimal intervention for control $a_2$ on gene CycA under the case only one control and the average reward of the optimal intervention for control on CycA using IMDP method with the other control $a_1$ on Rb ignored.</td>
</tr>
<tr>
<td>13</td>
<td>Illustrations of the average reward of the optimal intervention for control $a_1$ on gene Rb under the two controls case using IMDP method and the average reward of the optimal intervention for control on CycA under equilibrium</td>
</tr>
<tr>
<td>14</td>
<td>Illustrations of the average reward of the optimal intervention for control $a_2$ on CycA under the two controls case using IMDP method and the average reward of the optimal intervention for control on Rb under equilibrium</td>
</tr>
</tbody>
</table>
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSP</td>
<td>Genomic Signal Processing</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>T</td>
<td>Thymine</td>
</tr>
<tr>
<td>U</td>
<td>Uracil</td>
</tr>
<tr>
<td>RN</td>
<td>Regulatory Network</td>
</tr>
<tr>
<td>NCSG</td>
<td>Non-cooperative Stochastic Game</td>
</tr>
<tr>
<td>PBN</td>
<td>Probabilistic Boolean Network</td>
</tr>
<tr>
<td>NE</td>
<td>Nash Equilibrium</td>
</tr>
<tr>
<td>MDP</td>
<td>Markov Decision Process</td>
</tr>
<tr>
<td>GAP</td>
<td>Gene-activity Profile</td>
</tr>
<tr>
<td>IMDP</td>
<td>Independent Markov Decision Process</td>
</tr>
<tr>
<td>CE</td>
<td>Correlated Equilibrium</td>
</tr>
</tbody>
</table>
Genomic Signal Processing (GSP) is a discipline to study the processing of genomic signal. GSP studies a large collection of genomic sequence instead of individual gene. The aim of GSP is to integrate the theory and methods of signal processing with the global understanding of genomics. In GSP, one would find a proper representation of the genomic information and employ various signal processing methodologies such as detection, prediction, classification, control, etc.

In this thesis, we propose the concept of mapping equivalence theory for the numerical representation of symbolic data. We propose a framework for the analysis of different numerical mappings undergoing transformation by an analytic operator using Taylor’s expansion. Moreover, we emphasize the investigation of first- and second-order operators including the correlation function and Fourier transform. We also provide an analysis of the correlation between different numerical mappings of a symbolic sequence. In particular, we derive conditions for strong equivalence captured by perfect correlation among distinct mappings. We explore a relaxed similarity measure between distinct numerical mappings. Specifically, we provide conditions for weak equivalence which is characterized by preservation of the local extrema of the representation. We also introduce an abstract mapping model and extend the concept of equivalence to the generalized mapping model.

We extend the mapping equivalence theory for iterated operator. We provide a method for analyzing the consistency between different mappings under iterations of operator. We define
different concepts of mapping equivalence. We show the necessary and sufficient condition for consistency under iteration of affine operator. We present a few theoretical results for the equivalent mappings on the concept of Fatou and Julia Set. We give the definition of stability under iteration of operator and show the stability issue can be viewed as a special case of mapping equivalence. We also establish the connection of stability to Fatou and Julia set.

We propose the Non-cooperative stochastic game (NCSG) model for control of the genetic regulatory networks and formulate the intervention problem into solving the Nash equilibrium (NE). We show that the Markov decision process (MDP) is a special case of NCSG and the solving methods are provided. The definition of NE in this context has been proposed and the existences for both infinite and finite horizon cases have been proven. We provide a constructive method for solving the approximate NE.
CHAPTER 1

INTRODUCTION

1.1 Overview of Genomic Signal Processing and Regulatory Networks

Genomic Signal Processing (GSP) is a discipline to study the processing of genomic signal (1). GSP studies a large collection of genomic sequence instead of individual gene. The genomic signal is digital by nature, i.e. is represented as a collection of discrete and finite symbols. For example, the deoxyribonucleic acid (DNA) comprises of four different nucleotides, adenine (A), cytosine (C), guanine (G) and thymine (T). The ribonucleic acid (RNA) contains the same nucleotides as DNA with T substituted by uracil (U). In case of protein, there are 20 different amino acids.

The aim of GSP is to integrate the theory and methods of signal processing with the global understanding of genomics. In GSP, one would find a proper representation of the genomic information and employ various signal processing methodologies such as detection, prediction, classification, control, etc.

Each cell contains instructions necessary for its proper functioning, which are written in the form of DNA and must be replicated and handed down unchanged to the cell’s progeny when it divides. Coded in the DNA are instructions that direct the cell to divide, undergo apoptosis, or perform a variety of other functions. Although all these instructions are present in almost every cell of an organism, they are not active at all times. Control depends on complex interactions
between the products of the cell and those of its environment. As might be expected from a highly complex system that must be both efficient and survivable, control is highly distributed and redundant.

The key point is that cellular control, and its failure in disease, results from multivariate activity among cohorts of genes and their products. The control information as down and up regulation can be represented in a network, called regulatory network (RN). Essentially, cancer results when a cell divides uncontrollably and fails to undergo apoptosis. This can happen if damage to a cell’s DNA perpetually turns on the instructions to divide or permanently switches off the instructions to undergo apoptosis. The main purpose for study of regulatory network is to investigate the controlling method which will reverse the malignant status in a desired way.

1.2 Genomic Signal Representation and Dynamics

1.2.1 Genomic Signal Representation

In order to employ the various tools and methods in signal processing. Usually one need to find a representation of the genomic signals. Since the genomic signals are symbolic in natural, they are lack of algebraic structures. However, the majority portion of signal processing theory are built on the premise we have certain algebraic structure for the signal which is usually as a field. Therefore it is reasonable and important to investigate the effect of choosing different representations to the final analysis.

In general, symbolic information is represented as a sequence of symbols (possibly of infinite length) \( \{a_i\}_{i=0}^{X-1} \), where \( a_i \in \mathcal{A} \) and \( \mathcal{A} \) is a set of all possible symbols. For example, \( \mathcal{A} \) could be a collection of the 26-lowercase English letters, i.e. \( \mathcal{A} = \{a, b, \ldots, z\} \), or the four nucleotides
in a genomic sequence, i.e. $\mathcal{A} = \{A, T, G, C\}$. In statistical literature, symbolic data is usually called categorical data (2). The use of Markov chain models and hidden Markov models has been examined for time-domain analysis of genomic and proteomic data (3; 4; 5). However, we often seek to rely on frequency-domain analysis methods of symbolic signals. Unfortunately, symbolic sets do not generally possess an algebraic structure that allows us to define mathematical operations (e.g. group, ring, or field). In traditional signal processing, the set $\mathcal{A}$ corresponds to real- or complex-valued numbers, i.e. $\mathcal{A} = \mathbb{R}$ or $\mathbb{C}$, which form an algebraic field. However, attempts to define mathematical operations such as addition and multiplication on symbolic data has raised many questions about the meaning of the results obtained using such methods.

Several techniques exist which incorporate numerical and symbolic processing in an effective way to develop symbolic analysis systems (6). Software systems for symbolic computational algebra (e.g. Mathematica, Maple, etc.) represent a successful example of this approach. Such systems, however, are application-specific and difficult to realize for a broad class of symbolic signal processing applications. There are also various techniques for analyzing correlations, periodicities, etc. that do not require the aid of numerical symbol mappings. Among these techniques, the Mutual Information Function (MIF) (7) is one of the most important. The main advantage of these methods is that numerical mappings are not required. Moreover, it can be shown that methods such as MIF can capture any type of statistical dependence. The main disadvantages of these techniques, however, are that they generally provide less specific information than correlation analysis and they often suffer from a systematic overestimation of mutual information for finite sequences. Nevertheless, in order to extract the mathematical
and statistical information embedded in symbolic sequences, we wish to employ the powerful
analysis tools developed in traditional signal processing, e.g. Fourier transform, correlation
function, etc. We must therefore map the symbolic elements into numerical values. The re-
sulting numerical sequence should preserve the information embedded in the symbolic data.
Moreover, it should allow traditional signal processing techniques to extract the salient infor-
mation about the symbolic sequences from the corresponding numerical signals. For instance,
in DNA sequences, we have a finite alphabet associated with the four nucleotides in the genome,
i.e. $\mathcal{A} = \{A, T, G, C\}$. The mapping used for the representation of genomic data must preserve
the inherent structure of DNA sequences. In particular, if we choose a mapping such as: $A \mapsto 1,$
$T \mapsto 0,$ $G \mapsto -1,$ $C \mapsto 0,$ we would not preserve uniqueness since $T$ and $C$ are mapped to the
same value.

Numerous mappings have been proposed for the numerical representation of DNA sequences.
Buldyrev et al. (8) proposed various mapping rules for the representation of nucleotide sequences
into one-dimensional numerical sequences based on the purine-pyrimidine (RY) rule, hydrogen-
bond energy rule, etc. Li and Kaneko (9) and Voss (10) used the indicator sequence method,
which essentially maps the symbol to a standard basis of the 4-dimensional Euclidean space
$\mathbb{R}^4$. Berthelsen et al. (11) revised the method introduced in (8) by taking the molecular mass
and hydrophobicity into account in representation of genomic data. Silverman and Linsker
(12) relied on the simplex method, which maps the symbol to the vertices of a regular simplex.
Cristea and Anastassiou (13; 1) proposed the tetrahedral mapping, which maps the nucleotides
into corners of a tetrahedron. Stoffer et al. (14) introduced a mapping whose aim is to accentuate
the periodic features embedded in genomic sequences for stationary symbolic sequence analysis. Wang and Johnson (15) extended the method proposed by Stoffer et al. (14) for non-stationary sequence analysis. Rushid and Tuqan (16) proposed the Z-curve mapping, which is a unique 3-dimensional curve representation whose sequences are composed of binary values, i.e. 1 and −1. They also proposed a matrix-based framework to combine many widely used mapping strategies in genomic sequence analysis (17).

Each of the large number of numerical mappings used for the representation of genomic sequences can be justified for various applications. This raises several fundamental questions: What are the merits of each mapping used for the analysis of DNA sequences? How can we compare the results obtained from different numerical mappings? Indeed, it is impossible to determine which mapping is preferable. Furthermore, it is conceivable that distinct mappings could lead to contradictory conclusions. In fact, several contradictory results have arisen in the field of genomic sequence analysis. Most notably, the study of long-range correlations in coding and non-coding DNA sequences has been contested by several contradictory results (10; 18; 19). Investigation using a large DNA sequence database did not resolve this dispute; in fact, the controversy grew even further (20). Bouaynaya and Schonfeld (21; 22) shed light on this dilemma by demonstrating that a certain class of genomic sequences are inherently non-stationary and thus one of the reasons for the contradictory conclusions stems from the use of stationary time-series analysis tools. Moreover, they determined experimentally that the results obtained remained invariant over a large class of numerical mappings used for the representation of DNA sequences. Nonetheless, the experimental study conducted by Bouaynaya and Schonfeld
in (21; 22) cannot be used to ascertain with certainty whether the different numerical mappings used for representation of genomic sequences contributed to the contradictory findings reported in the literature (10; 18; 19).

1.2.2 Dynamics in Genomic Signal Processing

As in signal processing, the analysis is not commonly done in one step. Certain processing operations and many algorithms may be applied many times in processing the data. The consistency question rises again naturally in this situation. Another important question in this situation would be whether the analysis result is robust and stable to the perturbations of the data. Since noises and errors in acquisition of the biological data sequences are inevitable, the mapping equivalence or the stability issue under the iterations of operators can not be neglected. We need to establish a theory which can analyze the consistency and stability under the dynamics in signal processing.

The stability of dynamical system is a classical yet still active research area (23; 24; 25; 26). The Lyapunov stability theorem provides the sufficient condition for solution of differential equations (27). In signal processing, operators as de-noising filters, smoothing filters and certain algorithms may be applied iteratively in process of the data (28). Naturally, the important question that whether the analysis result is robust and stable to the perturbations of the data rises in this situation. Due to the noises and errors in acquisition of the data, the stability issue under the iterations of operators can not be neglected. However, because of the different assumptions of the model, many classical result of stability can not be applied easily in these cases.
The dynamical system in many of these situations could be modelled as the iteration of certain operator $\Phi : \mathbb{C}^N \to \mathbb{C}^N$. Therefore the stability is about whether a small enough perturbation of initial point will end up in a small change for the outcome under the iteration $\Phi^n$. The stability of this model is closely related to the complex dynamics (29; 30; 31). The concepts of Julia and Fatou set play an important role here.

1.2.3 Control in Regulatory Network

Cellular control and its failure in disease result from multivariate activity among cohorts of genes. For therapeutic purpose, it is important to model the cellular system as a regulatory network and propose techniques to control the network to prevent or stop the network entering into malicious states.

The probabilistic Boolean network (PBN) (32) is one the widely used model for genetic regulatory modeling. The purpose of intervention of the PBN is to reduce the probability to visit the unwanted states. The PBN can be modeled as a Markov chain model under certain assumptions (33). The perturbed gene can be seen as the external control of the Markov process. The Markov decision process (MDP) has been successfully applied to the intervention of PBN (34) and the optimal control policy can be solved by dynamic programming (35).

However, in many situation, the networks are overlapped in functioning. This means when we control different networks, they are intrinsically interacting with each other. In most cases, we are not aware or able to find this connections among these interactions. When we control the networks for different purposes, how should we model this situations and how should we control (36)?
1.3 **Summary of the Contributions**

The main contributions of my Ph.D. work include:

1. We propose the concept of mapping equivalence theory for the numerical representation of symbolic data. We propose a framework for the analysis of different numerical mappings undergoing transformation by an analytic operator using Taylor’s expansion. Moreover, we emphasize the investigation of first- and second-order operators including the correlation function and Fourier transform. We also provide an analysis of the correlation between different numerical mappings of a symbolic sequence. In particular, we derive conditions for strong equivalence captured by perfect correlation among distinct mappings. We explore a relaxed similarity measure between distinct numerical mappings. Specifically, we provide conditions for weak equivalence which is characterized by preservation of the local extrema of the representation. We also introduce an abstract mapping model and extend the concept of equivalence to the generalized mapping model.

2. We extend the mapping equivalence theory for iterated operator. We provide a method for analyzing the consistency between different mappings under iterations of operator. We define different concepts of mapping equivalence. We show the necessary and sufficient condition for consistency under iteration of affine operator. We present a few theoretical results for the equivalent mappings on the concept of Fatou and Julia Set. We give the definition of stability under iteration of operator and show the stability issue can be viewed as a special case of mapping equivalence. We also establish the connection of stability to Fatou and Julia set.
3. We propose the Non-cooperative stochastic game (NCSG) model for control of the genetic regulatory networks and formulate the intervention problem into solving the Nash equilibrium (NE). We show that the Markov decision process (MDP) is a special case of NCSG and the solving methods are provided. The definition of NE in this context has been proposed and the existences for both infinite and finite horizon cases have been proven. We provide a constructive method for solving the approximate NE.

Although GSP serves as a major application for these results, however, the theoretical results presented in 1 and 2 are applicable for any symbolic signal modeled by a discrete alphabet with a finite cardinality and independent of particular statistical properties such as stationarity, etc. The model and solution in 3 can also be applied to various other situations as communications and networks, economics, etc.
CHAPTER 2

MAPPING EQUIVALENCE FOR SYMBOLIC SEQUENCES

2.1 Introduction

As we mentioned in previous chapter, in GSP the signal we focus on is usually a sequences of nucleotides or amino acids. We need to process the symbolic data elements in order to extract useful information.

To ensure a clear understanding of the implications of the different choices used for numerical representation of symbolic data, we must develop a fundamental new approach that can be used to characterize the fundamental properties of numerical mappings. Specifically, it is essential that we establish a mapping equivalence theory for symbolic data that can be used to guarantee consistency among a class of numerical representations. With the aid of a mapping equivalence theory we could determine whether different mappings should yield compatible results, i.e. whether the mappings used for the analysis of the same data lead to consistent conclusions. Moreover, the theory can indicate when distinct mappings could lead to contradictory results and thus comparison of the corresponding conclusions is futile.

In this chapter, we provide a mapping equivalence theory for the numerical representation of symbolic data undergoing transformation by an operator (37; 38). We focus primarily on the mapping \( f : \mathcal{A} \to \mathbb{R}^n \) which maps the symbols to the \( n \)-dimensional Euclidean space. In Section 2.2, we first propose a framework for the analysis of different numerical mappings undergoing
transformation by an analytic operator using Taylor’s expansion. Moreover, we emphasize the investigation of first- and second-order operators including the correlation function and Fourier transform. These operators are widely used in signal processing and analysis and thus play an important role in this presentation. In Section 2.2.1, we provide an analysis of the correlation between different numerical mappings of a symbolic sequence. In particular, we derive conditions for strong equivalence captured by perfect correlation among distinct mappings. In Section 2.2.2, we explore a relaxed similarity measure between distinct numerical mappings. Specifically, we provide conditions for weak equivalence which is characterized by preservation of the local extrema of the representation. In Section 2.3, we introduce an abstract mapping model and extend the concept of equivalence to the generalized mapping model. In Section 2.4, we present experimental results which illustrate the significance of the proposed mapping equivalence theory in symbolic signal processing applications. Although the simulations are focused exclusively on analysis of genomic sequences, The results presented in this chapter, however, are applicable for any symbolic signal modeled by a discrete alphabet with a finite cardinality and independent of particular statistical properties such as stationarity, etc. Finally, we provide a brief summary and discussion of our results in Section 2.5.

2.2 Euclidean Mapping Equivalence for Symbolic Sequences

Given \( \{a_i\}_{i=0}^{N-1} \), where \( a_i \in \mathcal{A} \) and \( |\mathcal{A}| < \infty \), here \( |\cdot| \) means the cardinality of the set. \( f \) is a mapping from \( \mathcal{A} \) to \( \mathbb{R}^n \), i.e. \( f : a_i \mapsto x_i, \ x_i \in \mathbb{R}^n \). After the mapping we obtain a vector sequence \( \{x_i\}_{i=0}^{N-1} \). \( T : x_i \mapsto y_i \) is a transformation from \( \mathbb{R}^n \) to \( \mathbb{R}^m \). \( \Phi_l \) is an analytic operator.
on the numerical sequence and maps into \( \mathbb{R} \) parameterized by \( l \in \mathbb{R} \). We also assume that \( \Phi_l \in L^2(l) \). We classify the problems as in the following cases.

1. Given \( T \), determine the consistency between \( \Phi_l(\{x_i\}_{i=0}^{N-1}) \) and \( \Phi_l(\{T(x_i)\}_{i=0}^{N-1}) \). We need also figure out the largest class of operators which shows consistent results for two mappings under the given \( T \).

2. Given \( f \) and \( \Phi_l \), if \( f \) and \( T \circ f \) are consistent for any symbolic sequence \( \{a_i\}_{i=0}^{N-1} \). Find out the largest class of such transformation \( T \) which preserves the consistency. Also figure out the largest class of transformation \( T \) preserving the consistency for given mapping \( f \).

The consistency here means we require the results under two different mappings to be similar in certain extent. In general \( \Phi_l \) may not be linear. We will use Taylor’s expansion to expand the operator. We vectorize the vector sequence \( \{x_i\}_{i=0}^{N-1} \in \mathbb{R}^n \) to a large vector \( x \in \mathbb{R}^{Nn \times 1} \).

Consider the Taylor’s expansion of the analytic operator. \( \Phi_l : \mathbb{R}^{Nn \times 1} \rightarrow \mathbb{R} \). Without using the common scalar form representation of Taylor’s expansion (39), we shall present it in a concise form by using tensor product. First, we define the gradient operator \( \nabla \) as

\[
\nabla = \begin{pmatrix}
\frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \cdots & \frac{\partial}{\partial x_{Nn}}
\end{pmatrix}^T
\]

Then the Taylor’s expansion of \( \Phi_l \) at \( x_0 \) can be represented as the following form,
\[ \Phi_l = \sum_{i=0}^{\infty} \frac{1}{i!} (\nabla^i \Phi_l)(x_0) \times_i (X - x_0) \times_{i-1} (X - x_0) \times_{i-2} \cdots \times_1 (X - x_0) \] (2.2)

Where \( \times_i \) is the \( i \)th mode tensor product (40), and \( \nabla^i \) is the \( i \)th order gradient of \( \Phi_l \), which is defined as,

\[ \nabla^i \Phi_l = \Phi_l \times_1 \nabla \times_2 \nabla \times_3 \cdots \times_i \nabla \] (2.3)

Furthermore, \( \nabla^0 \Phi_l \) is defined as \( \Phi_l \). For one- and second-order terms, it is easy to check that it coincides with the well-known definition of Gradient \( \nabla \Phi_l(x) \) and Hessian \( \nabla^2 \Phi_l(x) \).

So we can rewrite the Taylor’s expansion at \( x_0 \) as,

\[ \Phi_l = \Phi_l(x_0) + \nabla \Phi_l(x_0)^T (x - x_0) + \]
\[ \frac{1}{2} (x - x_0)^T \nabla^2 \Phi_l(x_0) (x - x_0) \]
\[ + \sum_{i=3}^{\infty} \frac{1}{i!} (\nabla^i \Phi_l)(x_0) \times_i (x - x_0) \times_{i-1} (x - x_0) \times_{i-2} \cdots \times_1 (x - x_0) \] (2.4)

A metric or measure is needed for measuring the consistency. In general, there is no universal metric. Various operators may have different metrics for different purposes. In many cases, it
is a reasonable principle to require the results of two different mappings to be similar in some extent. In light of this principle, we propose the following two kinds of metrics.

### 2.2.1 Strong Equivalence: Perfect Correlation

We will use the correlation coefficient to characterize the consistency. First we provide the definition of the correlation coefficient $\rho$ used in this thesis.

**Definition 1** Given $\{a_i\}_{i=0}^{N-1}$, where $a_i \in \mathcal{A}$, $|\mathcal{A}| < \infty$. $f: a_i \mapsto x_i$, $x \in \mathbb{R}^n$, $T: x_i \mapsto y_i$ is a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, $\Phi_l$ is an operator on the numerical sequence. $m(\Phi_l) = \frac{1}{\mu(L)} \int_l \Phi_l d\mu$ is the mean value of the $\Phi_l$ in the space $L$ of parameter $l \in \mathbb{R}$. $\mu$ is a measure on $\mathbb{R}$. The correlation coefficient is defined as

$$
\rho = \frac{\int_l [\Phi_l(\{x_i\}_{i=0}^{N-1}) - m(\Phi_l(\{x_i\}_{i=0}^{N-1}))]}{\sqrt{\int_l [\Phi_l(\{x_i\}_{i=0}^{N-1}) - m(\Phi_l(\{x_i\}_{i=0}^{N-1}))]^2 d\mu}} \sqrt{\int_l [\Phi_l(\{T(x_i)\}_{i=0}^{N-1}) - m(\Phi_l(\{T(x_i)\}_{i=0}^{N-1}))]^2 d\mu}
$$

The use of abstract integration provides a unified framework for definition of the correlation coefficient. The measure $\mu$ can be chosen to be any Borel measure such as the Lebesgue-Stieltjes or counting measures depending on the properties of the operator. In practice, the measures we rely upon are mainly the counting measure and the Lebesgue measure.

It is well known that the correlation coefficient is between $[-1, 1]$ (41). The correlation coefficient can be used as a measure to characterize the similarity of two different mappings. For a given $T$, if $\rho = 1$, then we say the transformation $T$ is a *strongly equivalent* transformation of the map $f$ for an operator and $\Phi_l(\{T(x_i)\}_{i=0}^{N-1})$ is a *strong equivalence* of $\Phi_l(\{x_i\}_{i=0}^{N-1})$. When
the correlation coefficient is 1, it means the two mappings are the same up to a translation and scaling. This is the reason that it is called “strongly equivalent.” Unfortunately, there is no the universal equivalent transformation for arbitrary operator. However, because of the importance of second-order statistics, we shall emphasize on the second-order operators such as the correlation function. From now on we will focus on the transformation $T$ from $\mathbb{R}^n$ to $\mathbb{R}^n$.

For the case of mapping between Euclidean spaces with a different dimensions, we will present a detailed discussion in Section 2.2.4.

We first consider the correlation function. The correlation function of a sequence is defined as,

$$r_l = \frac{1}{N} \sum_{n=0}^{N-1} x^T(n)x(n-l)$$

(2.6)

Then if $\rho = 1$, we have the following theorem on the strongly equivalent transformation $T$.

**Theorem 1** For non-trivial operator and linear transformation $T$, the correlation coefficient $\rho = 1$ if and only if the transformation $T$ can be represented as $T(x_i) = \lambda Rx_i$, $R$ is an orthogonal matrix and $\lambda \in \mathbb{R}$.

**Proof 1** If $T(x_i) = \lambda Rx_i$, and $R$ is orthogonal. Then

$$r_l(T\{(x_i)_{i=0}^{N-1}\}) = \frac{1}{N} \sum_{n=0}^{N-1} \lambda x_i^T R^T R x(n-l)$$

$$= \frac{1}{N} \sum_{n=0}^{N-1} \lambda x_i^T x(n-l)$$

$$= \lambda r_l(\{x_i\}_{i=0}^{N-1})$$

(2.7)
Conversely, if $\rho = 1$ and $T(x_i) = Ax_i$. Then

$$r_l(T(\{x_i\}_{i=0}^{N-1})) = \frac{1}{N} \sum_{n=0}^{N-1} x_i A^T A x(n - l)$$

$$= \lambda' r_l(\{x_i\}_{i=0}^{N-1}) + c$$

$$= \lambda' \frac{1}{N} \sum_{n=0}^{N-1} x^T(n) x(n - l) + c$$

(2.8)

where $\lambda' \in \mathbb{R}$ and $c \in \mathbb{R}$ is a constant. Since the equality holds for any sequence and any $l$. So $A^T A = \lambda I$, then $A$ is orthogonal.

Actually, this property not only holds for correlation operator, but also for a larger class of operators. Consider the Taylor’s expansion of an operator $\Phi_l$. We would like first to introduce the definition of bounded linear operator and Riesz representation theorem (42). Then we will present a result for the first- and second-order bounded operators.

**Definition 2** Let $(X, \| \cdot \|)$ be a normed space. A operator $f : X \rightarrow \mathbb{R}$ is a bounded operator if $f$ is linear and there exists $C > 0$, such that $|f| \leq C\|x\|$.

The bounded operator can be thought as an analog of BIBO linear system in signal processing theory, which illustrates the good behaved operators. Furthermore if the space $X$ is a Hilbert space, we have the following theorem to characterize any linear bounded operator.

**Theorem 2** (Riesz Representation theorem for Hilbert space) $X$ is a Hilbert space, then for any linear bounded operator $\phi$, there exists a unique $y \in X$, such that $\phi(x) = < x, y >$. 
Note that $\mathbb{R}^n$ with the usual dot product is a Hilbert space. Therefore, Riesz Representation theorem for Hilbert space holds for $\mathbb{R}^n$. As before, we vectorize the vector sequence $\{x_i\}_{i=0}^{N-1} x_i \in \mathbb{R}^n$ to a large vector $x \in \mathbb{R}^{Nn \times 1}$. Then any linear transformation $T$ can be represented in the form

$$T = \begin{pmatrix}
A_{n \times n} & A_{n \times n} & \cdots & A_{n \times n} \\
A_{n \times n} & A_{n \times n} & \cdots & A_{n \times n} \\
& & \ddots & \cdots \\
& & \cdots & A_{n \times n}
\end{pmatrix}_{Nn \times Nn}$$

(2.9)

i.e. $y = Tx$.

Then we have the following theorem for equivalent transformation of first- and second-order operators.

**Theorem 3** Any non-trivial bounded linear operator does not have any non-trivial (scaled identity mapping) linear strongly equivalent transformation. If rotation is a strongly equivalent transformation for a bounded operator whose Taylor’s expansion does not have the third or higher-order terms, then its Taylor’s expansion can not have first-order term and the Hessian $\nabla^2 \Phi_l(x)$ must have the form

$$\nabla^2 \Phi_l(x) = \begin{pmatrix}
k_{11}I_{n \times n} & k_{12}I_{n \times n} & \cdots & k_{1N}I_{n \times n} \\
k_{21}I_{n \times n} & k_{22}I_{n \times n} & \cdots & k_{2N}I_{n \times n} \\
& & \ddots & \cdots \\
k_{N1}I_{n \times n} & k_{N2}I_{n \times n} & \cdots & k_{NN}I_{n \times n}
\end{pmatrix}$$

(2.10)
where \( k_{ij} \in \mathbb{R} \) and \( k_{ij} = k_{ji}, \forall i \neq j \).

**Proof 2** Notice that \( \mathbb{R}^n \) with inner product \( \langle x, y \rangle = x^T y \) is a Hilbert space. So for any linear bounded operator, \( \exists y \in \mathbb{R}^n \), such that \( \Phi_1(x) = x^T y \). So \( \Phi_1(T(x)) = T^T x^T y \). \( T \) is strongly equivalent, therefore \( \Phi_1(T(x)) = \lambda \Phi_1(x) + c \) for some \( \lambda \) and \( c \). Then we have \( T = \lambda I_{Nn \times Nn} \), i.e. \( T \) is a trivial scaled identity transform. This finishes the proof of the first claim. We claim if a non-trivial operator \( \Phi_1(x) \) whose Taylor’s expansion has no terms of order higher than or equal to three has a non-trivial linear strongly equivalent transformation, then it must only have the second order term and the constant term. We can always scale or add constant for the transformation to get a strongly equivalent result after transformation. So without loss of generality, we assume the result after the transformation is exactly the same as the previous one, i.e. If \( \Phi_1(x) = \frac{1}{2} x^T \nabla^2 \Phi_1(0) x \) and \( \Phi_1(T(x)) = \Phi_1(x) \), then we have

\[
\frac{1}{2} x^T \nabla^2 \Phi_1(0) x = \frac{1}{2} x^T T^T (\nabla^2 \Phi_1(0)) T x \tag{2.11}
\]

this equality holds for any \( x \in \mathbb{R}^{Nn \times 1} \). Therefore we have \( T^T (\nabla^2 \Phi_1(0)) T = \nabla^2 \Phi_1(0) \). \( T \) is a rotation, i.e.

\[
T = \left( \begin{array}{c}
R_{n \times n} \\
& R_{n \times n} \\
& & \ddots \\
& & & R_{n \times n}
\end{array} \right)_{Nn \times Nn} \tag{2.12}
\]
Since \( R^T R = I \), we have

\[
(\nabla^2 \Phi_l(0))^T = T \nabla^2 \Phi_l(0)
\]  

(2.13)

Because \( R^T R = RR^T = I \), \( T^T T = TT^T \), therefore \( T \) is normal, \( T \) is unitarily diagonalizable (43). \( R \) is also normal. Therefore \( \exists V \) unitary, such that \( R = V^H \Lambda' V \), where \( \Lambda' \) is a diagonal matrix. Since \( R \) is real orthogonal, the eigenvalues of \( R \) are on the unit sphere \( S^1 \). Without loss of generality, we assume \( R \) has two eigenvalues, 1 and \( \mu \). Let the algebraic multiplicity of 1 be \( i \), then the algebraic multiplicity of \( \mu \) is \( n - i \). So we have

\[
\Lambda' = \begin{pmatrix}
I_{i \times i} \\
\mu I_{(n-i) \times (n-i)}
\end{pmatrix}
\]  

(2.14)

Let \( U = \begin{pmatrix}
V \\
\cdots \\
V
\end{pmatrix} \), \( \Lambda = \begin{pmatrix}
\Lambda' \\
\cdots \\
\Lambda'
\end{pmatrix} \). From (Equation 2.13), we have \( (\nabla^2 \Phi_l(0))^T U^H \Lambda U = U^H \Lambda U \nabla^2 \Phi_l(0) \). Therefore we have

\[
U(\nabla^2 \Phi_l(0)) U^H \Lambda = \Lambda_U \nabla^2 \Phi_l(0) U^H
\]  

(2.15)

Let \( \tilde{X} = U(\nabla^2 \Phi_l(0)) U^H \). We have \( \tilde{X} \Lambda = \Lambda \tilde{X} \). By using the Jordan canonical form (44, Chapter VIII), we have that all \( \tilde{X} \) which commutes with \( \Lambda \) must have the form as follow:
Every non-zero submatrix in $\tilde{X}$ is an arbitrary upper triangular submatrix which has identical diagonal entries. All submatrices $A_{kl}$ and $A'_{kl}$ have size $i \times i$, all $B_{kl}$ and $B'_{kl}$ have size $(n-i) \times (n-i)$. Thus all $\nabla^2 \Phi_l(0)$ satisfies (Equation 2.13) are of the form $\nabla^2 \Phi_l(0) = U^H \tilde{X} U$. However, since the $\Phi_l(x)$ is analytic, the Hessian must be symmetric. Therefore all the submatrices of $\tilde{X}$, $A_{kl}$ and $B_{kl}$ have the form $k_{kl}I$ and $A'_{kl}$ and $B'_{kl}$ have also the form $k_{kl}I$. Notice that this is for a given rotation. Since (Equation 2.13) holds for any given rotation, we have that

$$\nabla^2 \Phi_l(0) \in Y = \bigcap_{U \text{ Unitary}} \{U^H \tilde{X} U\} \quad (2.17)$$
The rotation is arbitrary, (Equation 2.14) should hold for any $i = 0, \ldots, n$. Claim that the $N \times N$ principal matrices

$$
\begin{pmatrix}
A(i)_{ij} \\
B_{(i+1)j}
\end{pmatrix}
\quad\text{and}\quad
\begin{pmatrix}
A'(i)_{ij} \\
B'_{(i+1)j}
\end{pmatrix}
$$
in $\tilde{X}$ must satisfy:

$$
\begin{pmatrix}
A(i)_{ij} \\
B_{(i+1)j}
\end{pmatrix} = k_{ij}I_{n \times n} \tag{2.18}
$$

$$
\begin{pmatrix}
A'(i)_{ij} \\
B'_{(i+1)j}
\end{pmatrix} = k'_{ij}I_{n \times n} \tag{2.19}
$$

where $i = 1, 3, 5, \ldots, (2N - 1)$ and $j = 1, \ldots, M$. Because we know that

$$
\begin{pmatrix}
A(i)_{ij} \\
B_{(i+1)j}
\end{pmatrix} = \begin{pmatrix}
k_{ij}I_{i \times i} \\
k_{(i+1)j}I_{(n-i) \times (n-i)}
\end{pmatrix} \tag{2.20}
$$

but if $k_{ij} \neq k_{(i+1)j}$, then this implies $\{U^H \tilde{X}U\} \cap \{U^H \tilde{X}'U\} = \emptyset$, where (Equation 2.14) for $\tilde{X}'$ is of the form

$$
\Lambda' = \begin{pmatrix}
I_{(i+1) \times (i+1)} \\
\mu I_{(n-i-1) \times (n-i-1)}
\end{pmatrix} \tag{2.21}
$$

Since we know $Y$ is not empty, we get a contradiction here. Therefore (Equation 2.18) and (Equation 2.19) hold. If we choose $U = I$, we have $Y \subset \tilde{X}$, where for $\tilde{X}$, (Equation 2.18)
and (Equation 2.19) hold. It’s straightforward to check that such \( \tilde{X} \) is commutable with \( T \). Therefore \( Y = \tilde{X} \). Finally we show that

\[
\nabla^2 \Phi_l(0) = \begin{pmatrix}
k_{11}I_{n \times n} & k_{12}I_{n \times n} & \cdots & k_{1N}I_{n \times n} \\
k_{21}I_{n \times n} & k_{22}I_{n \times n} & \cdots & k_{2N}I_{n \times n} \\
\vdots & \vdots & \ddots & \vdots \\
k_{N1}I_{n \times n} & k_{N2}I_{n \times n} & \cdots & k_{NN}I_{n \times n}
\end{pmatrix}
\]

(2.22)

where \( k_{ij} \in \mathbb{R} \) and \( k_{ij} = k_{ji}, \forall i \neq j \). If we expand at any other point \( x_0 \), then

\[
\frac{1}{2}(x-x_0)^T \nabla^2 \Phi_l(x)(x-x_0) = \frac{1}{2}x^T \nabla^2 \Phi_l(x_0)x - \frac{1}{2}x_0^T \nabla^2 \Phi_l(x_0)x - \frac{1}{2}x^T \nabla^2 \Phi_l(x_0)x_0 + \frac{1}{2}x_0^T \nabla^2 \Phi_l(x_0)x_0
\]

(2.23)

The only second order term is \( \frac{1}{2}x^T \nabla^2 \Phi_l(x_0)x \). Repeat the previous argument, we will have (Equation 2.10).

For Fourier transform, in many situations, we focus exclusively on the modulus of the transform of symbolic data, i.e. we discard the phase information. Since the module of continuous-time Fourier transform is invariant under rotation, it is tempting to conclude that rotation is an equivalent transformation for the Fourier transform. However, the widely used form of the Fourier transform used in much of the literature devoted to DNA sequence analysis (15) is dif-
ferent from the classical multi-dimensional Fourier transform. Fortunately, we are able to show
that rotation still yields an equivalent transformation. We first define the Fourier transform as:

$$\hat{f}_m = \frac{1}{N^2} ||X L_F||^2_2$$ (2.24)

where $X$ is a $n \times N$ matrix, whose $i^{th}$ column is $x_i$. $L_F$ is the frequency vector, i.e.

$$L_F = \begin{pmatrix} e^{-2\pi jm_0/N} & e^{-2\pi jm_1/N} & \cdots & e^{-2\pi jm_{(N-1)/N}/N} \end{pmatrix}^T$$ (2.25)

If we vectorize $X$ to $x \in \mathbb{R}^{Nn \times 1}$ as before, $\hat{f}_m$ can also be represented as $\hat{f}_m = \frac{1}{N^2} ||Ax||^2_2$, where

$$A = \begin{pmatrix} e^{-2\pi jm_0/N} I_{n \times n} & \cdots & e^{-2\pi jm_{(N-1)/N}/N} I_{n \times n} \end{pmatrix}_{n \times N}$$ (2.26)

Notice that $\hat{f}_m = \frac{1}{N^2} (x^H A^H A x)$, which is a second-order operator and $A^H A$ is of the form,

$$A^H A = \begin{pmatrix} \bar{w}_0 w_0 I_{n \times n} & \bar{w}_0 w_1 I_{n \times n} & \cdots & \bar{w}_0 w_N I_{n \times n} \\
\vdots & \vdots & \ddots & \vdots \\
\bar{w}_N w_0 I_{n \times n} & \bar{w}_N w_1 I_{n \times n} & \cdots & \bar{w}_N w_N I_{n \times n} \end{pmatrix}$$ (2.27)

where $w_i = e^{-2\pi jm_i/N}$. By theorem 3, rotation is a strongly equivalent transformation for Fourier
transform.
2.2.2 Weak Equivalence: Preservation of Local Extrema

In the previous section, we employed the correlation coefficient as a metric to characterize the similarity for an operator under transformation. However, as we can see, the strong equivalence basically requires the result to be “exactly” the same. While in many situations, we do not focus on whether or not the result under two mapping strategies are exactly the same, i.e. the true numerical value of the result, but the relative relation or the relative trend of the result. For example, when we use the correlation function, in many cases, we only care where the peak and valley points are located and the changing trends, which are used to determine the periodicity structure of certain patterns. In these cases, what we really need is to preserve the local extremums and local trend under the a transformation. So we first give the definition of Local Minimum and Maximum Preserving Similarity or in this thesis what we call weakly equivalent.

**Definition 3** Given \( \{a_i\}_{i=0}^{N-1} \), where \( a_i \in A, |A| < \infty \). \( f : a_i \mapsto x_i, x \in \mathbb{R}^n, T : x_i \mapsto y_i \) is a transformation from \( \mathbb{R}^n \) to \( \mathbb{R}^m \), \( \Phi \) is an operator on the numerical sequence. We say \( T \) is weakly equivalent, if for every \( x \), which is a local minimal or maximals for \( \Phi_x(\{x_i\}_{i=0}^{N-1}) \) then \( \Phi_x(\{T(x_i)\}_{i=0}^{N-1}) \) is also a local minimal or maximal respectively.

A few easy observations and results follow. By definition strong equivalence implies weak equivalence. Moreover, we have the following propositions to determine weak equivalence.
Proposition 1 If $\Phi_l$ is twice differentiable with respect to $l$, then $T$ is weakly equivalent, if for any $l$, where $\frac{\partial \Phi_l(\{x_i\}_{i=0}^{N-1})}{\partial l} = 0$, the following conditions hold

$$\frac{\partial \Phi_l(\{T(x_i)\}_{i=0}^{N-1})}{\partial l} = 0$$ (2.28)

and

$$\frac{\partial^2 \Phi_l(\{x_i\}_{i=0}^{N-1})}{\partial l^2} \cdot \frac{\partial^2 \Phi_l(\{T(x_i)\}_{i=0}^{N-1})}{\partial l^2} \geq 0$$ (2.29)

Proof 3 if $l$ is a local maximal or local minimal, then $\frac{\partial \Phi_l(\{x_i\}_{i=0}^{N-1})}{\partial l} = 0$ and $\frac{\partial^2 \Phi_l(\{x_i\}_{i=0}^{N-1})}{\partial l^2} \leq 0$ or $\frac{\partial^2 \Phi_l(\{x_i\}_{i=0}^{N-1})}{\partial l^2} \geq 0$. By the definition of weak equivalence, (Equation 2.28) and (Equation 2.29) follow.

If $l \in \mathbb{Z}$, Then we have the following criterion to determine weak equivalence.

Proposition 2 $T$ is weakly equivalent for an operator $\Phi_l$, where $l \in \mathbb{Z}$, if for any $l$, the following condition holds

$$(\Phi_l(\{x_i\}_{i=0}^{N-1}) - \Phi_{l-1}(\{x_i\}_{i=0}^{N-1})) \cdot (\Phi_l(\{T(x_i)\}_{i=0}^{N-1}) - \Phi_{l-1}(\{T(x_i)\}_{i=0}^{N-1})) \geq 0$$ (2.30)

Proof 4 Without loss of generality, we assume $l$ is a local maximal for $\Phi_l(\{x_i\}_{i=0}^{N-1})$, then

$\Phi_l(\{x_i\}_{i=0}^{N-1}) \geq \Phi_{l-1}(\{x_i\}_{i=0}^{N-1})$ and $\Phi_l(\{x_i\}_{i=0}^{N-1}) \geq \Phi_{l+1}(\{x_i\}_{i=0}^{N-1})$. If (Equation 2.30) holds,
we have \( \Phi_l(\{T(x_i)\}_{i=0}^{N-1}) \geq \Phi_{l-1}(\{T(x_i)\}_{i=0}^{N-1}) \) and \( \Phi_l(\{T(x_i)\}_{i=0}^{N-1}) \geq \Phi_{l+1}(\{T(x_i)\}_{i=0}^{N-1}). \)

Thus \( l \) is also a local maximal for \( \Phi_l(\{T(x_i)\}_{i=0}^{N-1}) \).

As the importance of second-order statistics, specially we would like to investigate the weakly equivalent transformation for the correlation function. We first introduce a lemma.

**Lemma 3** If the transformation \( T: \mathbb{R}^n \rightarrow \mathbb{R}^n \) is an inner-product preserving isometry, i.e. \( x^T y = T(x)^T T(y), \forall x,y \in \mathbb{R}^n \), then \( T(x) = Rx \), where \( R \) is an orthogonal matrix. Hence \( T \) is a bijective isometry.

**Proof 5** First let \( x = y \), we have \( ||x||_2 = ||T(x)||_2 \), i.e. \( T \) preserves the Euclidean norm. Since \( T(x_0) \) is on the ball \( ||x||_2 = ||x_0||_2 \), we have \( T(x) = R(x)x \), where \( R(x) \) is an orthogonal matrix function. Let

\[
R(x) = ( u_1(x) \quad u_2(x) \quad \ldots \quad u(x))
\]

(2.31)

, where \( \{u_i(x)\} \) is orthonormal. Furthermore, let \( x = (1,0,0,\ldots,0)^T \) and \( y = (y_1,0,0,\ldots,0)^T. \) then we have

\[
x^Ty = T(x)^T T(y)
\]

(2.32)

\[
y_1 = xR(x)^TR(y)y = y_1u_1^T((1,0,0,\ldots,0)^Tu_1(y)
\]

(2.33)
Therefore $u_1^T((1, 0, 0, \ldots, 0)^T)u_1(y) = 1$. By Cauchy-Schwartz Inequality, we have

$$1 = u_1^T((1, 0, 0, \ldots, 0)^T)u_1(y) \leq ||u_1((1, 0, 0, \ldots, 0)^T)||_2 \cdot ||u_1(y)||_2 = 1$$

(2.34)

The equality holds if and only if $u_1(y) = \lambda u_1((1, 0, 0, \ldots, 0)^T)$, but $||u_1(y)||_2 = 1$, thus $\lambda = 1$. Therefore

$u_1(y) = u_1((1, 0, 0, \ldots, 0)^T), \forall y \in \mathbb{R}^n$. By the same arguments we can show $u_i(y) = u_i(e_i), \forall i = 1, \ldots n$, where $e_i$ is the standard basis of $\mathbb{R}^n$. So $R(x) = R$ is a constant orthogonal matrix, thus $T(x) = Rx$. This also shows $T$ is a bijective isometry.

For correlation function, we have the following theorem showing that generally speaking, rotation can be thought as the "only" weakly equivalent transformation.

**Theorem 4** For a fix length sequence, any transformation which only brings small enough changes to the inner product value under previous mapping will be a weakly equivalent transformation for correlation function. However, if the length goes to infinity, then rotation (or scaled rotation) is the only weakly equivalent transformation for correlation function.

**Proof 6** Consider the vector sequence $\{x_i\}_{i=0}^{N-1}$. The correlation function is

$$r_l(\{x_i\}_{i=0}^{N-1}) = \frac{1}{N} \sum_{n=0}^{N-1} x^T(n)x(n - l).$$

Then we have

$$r_l - r_{l-1} = \frac{1}{N} \sum_{n=0}^{N-1} x^T(n)(x(n - l) - x(n - l + 1))$$

(2.35)
After the transformation, we have the correlation function

\[ r'_l = r_l(T(\{x_i\}_{i=0}^{N-1})) = \frac{1}{N} \sum_{n=0}^{N-1} T(x(n))^T T(x(n-l)) \]

and

\[ r'_l - r'_{l-1} = \frac{1}{N} \sum_{n=0}^{N-1} T(x(n))^T (T(x(n-l)) - T(x(n-l+1))). \]

By Proposition 2, \( T \) is weakly equivalent if \( r'_l - r'_{l-1} \) has the same sign as \( r_l - r_{l-1} \).

Consider the alphabet \( \mathcal{A}' = \mathcal{A} \cup \{a-b|a, b \in \mathcal{A}\} \), which means we consider the symbol “\( a-b \)” as a new symbol and we extend the mapping \( f \) on the newly added symbols as \( f: (a-b) \rightarrow (f(a) - f(b)) \), which also extends the transformation \( T \) for \( f(a-b) \) be \( T(f(a)) - T(f(b)) \). Thus finding the weakly equivalent transformation is same to find the \( T \) which preserve the sign at each \( l \) of cross correlation function \( R_l \) for the sequence \( \{x_i\}_{i=0}^{N-1} \) and \( \{x_i - x_{i+1}\}_{i=0}^{N-1} \).

Let \( \mathcal{F} = (a_1 \ a_2 \ \ldots \ a_A)^T \), where \( a_i = f(x)^T f(y) \), \( (x, y) \) or \( (y, x) \in \mathcal{A} \times \mathcal{A}' \) and \( A = |\mathcal{A} \times \mathcal{A}'| \). Also let \( N_l = (C_1 \ C_2 \ \ldots \ C_A)^T \), where \( C_i \) is the counting number for the pair \((x, y)\) corresponding to \( a_i \) which appears in the cross correlation function \( R_l \). Therefore we have

\[ R_l = \frac{1}{N} \mathcal{F}^T N_l \quad (2.36) \]

Define \( T(\mathcal{F}) = (b_1 \ b_2 \ \ldots \ b_A)^T \), where \( b_i = T(f(x))^T T(f(y)) \), where \((x, y)\) corresponds to \((x, y)\) in the \( a_i \) in \( \mathcal{F} \). Notice that \( N_l \) will not change since it is determined by the given sequence. After the transformation, the cross correlation functions becomes \( R'_l = \frac{1}{N} T(\mathcal{F})^T N_l \). We need \( R_l \) and \( R'_l \) have the same sign for all \( l \). Notice that \( \sum_{i=1}^{A} C_i = N \). So every \( N_l \) for a given sequence corresponds to a point on the hyperplane \( \sum_{i=1}^{A} x_i = N \) in \( \mathbb{R}^A \). If \( T \) preserves signs for all \( l \), then for each \( N_l \), \( T(\mathcal{F}) \) should reside in the same half plane of \( \mathcal{F} \). Because
Figure 1. Illustration for $N_l$ and $T(\mathcal{F})$. $T(\mathcal{F})$ should reside in the convex cone as the shaded area in the figure.

$T(\mathcal{F})^T N_l$ should have the same sign of $\mathcal{F}^T N_l$. In general, the sign will not all be positive or negative, since that will mean the correlation function is monotonic which in general is not valid for all sequences. Consider all possible symbol sequence of length $N$ and $l \in \mathbb{Z}$. Then

$$T(\mathcal{F}) \in \bigcap_{\text{all sequences of length } N} \{ x \mid (x^T N_l) \cdot (\mathcal{F}^T N_l) \geq 0 \}$$  \hspace{1cm} (2.37)

This implies $T$ should reside in the intersection of all half plane determined by all sequences of length $N$. Each half plane is a convex cone, therefore the intersection is still a convex cone as illustrated in Figure 1. Since we have finitely many point on the hyperplane $\sum_{i=1}^{A} x_i = N$. All $N_l$ reside in the first quadrant. We can always construct two hyperplanes whose intersection is in one quadrant. Thus the intersection will be a small convex cone in a quadrant. If $T(\mathcal{F})$
is in that convex cone, then the sign is all preserved, which means \( T \) is a weakly equivalent transformation. This proves the first claim.

However, if we let \( N \) go to infinity, first notice that the intersection will not be empty set, since \( \mathcal{F} \) is always in the intersection. But the points \( N_i \) become dense in the first quadrant. We can have a sequence of \( N_i \) such that \( N_i^T \mathcal{F} \to 0^+ \) and \( N_i' \) with \( N_i' \mathcal{F} \to 0^- \). Therefore the intersection will be squeezed to the line \( \lambda \mathcal{F}, \lambda \geq 0 \). i.e. The \( T(\mathcal{F}) = \lambda \mathcal{F} \). If the scaling \( \lambda = 1 \), then \( T(f(x))^T T(f(y)) = f(x)^T f(y), \forall f \). Since the mapping is arbitrary, which means it should hold for any \( x \) and \( y \). By lemma 3, \( T \) is a rotation or a scaled rotation.

Conversely, by theorem 1, we have rotation or scaled rotation is a strongly equivalent transformation. Thus it follows it is also a weakly equivalent transformation.

From this theorem, we can see that rotation is essentially the only weakly equivalent transformation for correlation function. We can expand the class of operator having this property by combining theorem 4 and theorem 3 with some technical conditions, then we have the following corollary.

**Corollary 4** If a second order operator whose Hessian is of the form as (Equation 2.10) and all \( k_i \) have the same sign and \( 0 < \sum k_i < \infty \), then rotation is essentially the only transformation which is both strongly and weakly equivalent.

**Proof 7** If the Hessian has the form above, then

\[
\Phi_l = \frac{1}{2} \text{sgn}(k_1) \sum_{n=0}^{N-1} |k_n| x(n)^T x(n - l)
\]

(2.38)
Where

\[
\text{sgn}(x) = \begin{cases} 
1 & x \geq 0 \\
-1 & x < 0 
\end{cases}
\]

Notice that the proof for correlation function follows here except we shall use \(N'_l = (\zeta_1 C_1 \ z_2 C_2 \ \ldots \ \zeta_A C_A)^T\) instead of \(N_l\), where \(\zeta_i\) are all positive and

\[
\zeta_i = \sum_{j \text{ all terms having same pattern}} k_j
\]

If the length goes to infinity, \(N_l\) becomes dense, since \(0 < \sum k_i < \infty\), then \(N'_l\) is in some non-degenerated bounded set, which is still dense and resides in the first quadrant. Then the argument above still valid. Thus rotation is essentially the only weakly equivalent for this kind of operator. By theorem 1, rotation is also the strongly equivalent transformation.

### 2.2.3 Mapping Between 1D Euclidean Spaces

The case in which the transformation \(T: \mathbb{R}^n \to \mathbb{R}^n\) is limited to \(n = 1\) is particularly common in the literature. Moreover, this case stands out and deserves special attention since the only possible rotation on \(\mathbb{R}\) is obtained by scaling and interchanging the mapping values. From Theorems 1 and 4, we observe that if the mappings cannot be obtained by scaling and interchanging the mapping values, the correlation and Fourier analysis results obtained using these mappings are neither strongly nor weakly equivalent. Therefore, in the 1D case, the equivalent mapping class under a given operator becomes fairly limited. In general, distinct mappings will usually lead to inconsistent correlation and Fourier analysis results. Another
interesting fact about 1D mappings that can be derived from our previous results is that if the mapping is binary (i.e. the range of the mapping can only takes two distinct values), then we observe that the correlation and Fourier analysis under any two such binary mappings are always consistent since we can always obtain one of the mappings by scaling and interchanging the mapping values of the other mapping.

2.2.4 Mapping Between Euclidean Spaces of Different Dimensions

In the previous sections, we focused primarily on the transformation \( T : \mathbb{R}^n \rightarrow \mathbb{R}^m \), where \( n = m \). In this section, we will present a brief discussion of the case where \( m \neq n \). If \( m > n \), which means \( T \) will transform the vector into a larger dimensional Euclidean space. However, since there is a natural embedding for \( \mathbb{R}^n \) into \( \mathbb{R}^m \), we can always think the transformation as \( T' : \mathbb{R}^m \rightarrow \mathbb{R}^m \). For second-order operators which are shown equivalent under rotation, we still have the same results in this situation, except the rotation matrix here means a matrix have orthonormal columns.

For the case \( m < n \), we can also think as \( \mathbb{R}^m \) is embedded inside \( \mathbb{R}^n \) by the transform

\[
y = \begin{pmatrix}
    I_{m \times m} \\
    0_{(n-m) \times m}
\end{pmatrix}_{n \times m} x
\]

\hspace{1cm} (2.39)

, where \( x \in \mathbb{R}^m \), \( y \in \mathbb{R}^n \). Then we only need to research on the new transformation \( T' : \mathbb{R}^n \rightarrow \mathbb{R}^n \). However, in this case, we can see that we actually project the higher dimensional subspace into a lower dimensional space, rotation here in general is not an equivalent transformation anymore.
Intuitively, because of the projection, we lose information projected on \((n - m)\) dimensions. Therefore rotation is no longer an equivalent transformation.

2.3 Abstract Mapping Equivalence for Symbolic Sequences

2.3.1 Abstract Mapping Model and Examples

In previous sections, we mainly focused on properties of mappings, which map symbols into vector space. However, it is not necessary to restrict to the vector space. Many classical concepts in numerical signal processing can be extended to various algebraic structures. For example, the Fourier transform and Wavelet transform can be defined on group, ring and finite-field \((45; 46; 47)\). In this section, we introduce the generalized mapping to arbitrary semi-ring, ring or algebra structure. We shall also extend the notion of equivalence defined in previous section.

For given finite alphabet \(A\), we define \(F(A)\) as collection of all the symbolic sequences and define the binary operation as concatenating two symbolic sequences. It can be shown \(F(A)\) is a free semi-group \((48)\). \(R\) is any semi-ring. Let \(R\) to be collection of all maps from \(F(A)\) to \(R\). For any \(f \in R\), we denote it as the formal series,

\[
f = \sum_{u \in F(A)} f(u)u
\]  

\(2.40\)
we define two operations $+$ and $\cdot$ on $\mathcal{R}$ as

$$f + g = \sum_{u \in F(\mathcal{A})} (f(u) + g(u))u$$  \hspace{1cm} (2.41)

$$ (f \cdot g)(s) = \sum_{uv=s} (f(u)g(v))s$$  \hspace{1cm} (2.42)

With these two binary operations, we have the following proposition to show that we construct a new algebraic structure on $\mathcal{R}$

**Proposition 5**  $R$ is a semi-ring or ring, then $(\mathcal{R}, + , \cdot)$ forms a semi-ring or ring respectively.

**Proof 8**  By the definition of addition, we can see that if $(\mathcal{R}, +)$ is a commutative monoid or abelian group, then $(\mathcal{R}, +)$ has the same property correspondingly. Therefore it’s enough to show to show $(\mathcal{R}, \cdot)$ is a semi-group, i.e. we need to show the multiplication is associative.

$\forall f, g, h \in \mathcal{R}$

$$ (fg)h(s) = \sum_{xw=s} (fg(x))h(w) = \sum_{xw=s} \sum_{uv=x} (f(u)g(v))h(w)$$

$$ = \sum_{u(vw)=s} f(u)(g(v)h(w)) = f(gh)(s) \hspace{1cm} (2.43)$$

Therefore the multiplication is associative. We proved the proposition.
The ring $\mathcal{R}$ is called as the semi-group ring of $F(A)$ with coefficients in $R$. Furthermore, if $R$ is a left-$R'$ module for some ring $R'$. We can define for any $r \in R'$,

$$ rf = \sum_{u \in F(A)} rf(u)u $$

then $\mathcal{R}$ is the left-$R'$ algebra. The $\mathcal{R}$ may be interpreted as the generalized filter space, while the $F(A)$ is as the signal space. The multiplication can been thought as the extension of discrete convolution. If we let $R$ and $R'$ to be $\mathbb{R}$, $A = \{0, 1, 2, \ldots\}$, the multiplication degenerates to classical convolution. The symbol sequence is mapped into a numerical sequence.

Another example of the abstract mapping model is the probability model. Consider all the outcomes of the words in $F(A)$. Denote the outcome space as $\Omega$. $\omega$ is a $\sigma$-algebra on $\Omega$ and $\mathbb{P}$ is a probability measure on $\sigma$. Notice that two set-operations on $\sigma$, $\cap$ and $\cup$ is analog of $\cdot$ and $\cdot$. $\emptyset$ and $\Omega$ can be seen as 0 and 1 respectively. Therefore $R = (\omega, \cup, \cap)$ forms a semi-ring. The probability measure mapping $P$ here is interpreted as a semi-ring mapping from $\mathcal{R}$ to the semi-group ring of $F(A)$ with coefficients in $\mathbb{R}$, which is defined as,

$$ P(f) = \sum_{u \in F(A)} \mathbb{P}(f(u))u $$

The probability operations then can be realized by algebraic operations on $\mathcal{R}$ and the corresponding probability measure values are obtained after the mapping $P$. 
2.3.2 Abstract Mapping Equivalence

For generalized mapping, the equivalence problem is still worth for investigating. However, in the situation, it becomes much more difficult than in a $\mathbb{R}$-vector space. The $R$ in general does not possess any meaningful ordering. Therefore the definition of equivalence turns out to be limited for specific application. Nevertheless, as we mentioned before, in most cases, it is reasonable to require the result to be similar in certain extent. From now on, we assume $R$ is an integral domain with unity 1. We introduce the following definition for abstract equivalence of a generalized mapping.

**Definition 4** For any $f, g \in R$, $f$ and $g$ are abstractly equivalent, if the ideals they generated are the same, i.e. $(f) = (g)$.

The next proposition shows the intuition and legitimacy of this definition.

**Proposition 6** $(f) = (g)$ if and only if $f = ug$, where $u$ has multiplicative inverse.

**Proof 9** If $(f) = (g)$, then $f = u_1 g$ and $g = u_2 f$ for some $u_1, u_2 \in R$. We have

$$f = u_1 u_2 f$$

$$(1 - u_1 u_2)f = 0 \quad (2.46)$$

Since $R$ is integral domain, we have $u_1 u_2 = 1$. $u_1$ and $u_2$ are units.

Conversely, if $f = ug$, then $g = u^{-1} f$. We have $(f) \subseteq (g)$ and $(g) \subseteq (f)$, therefore $(f) = (g)$. 
A loose interpretation of Proposition 6 implies that abstractly equivalent mappings only differ by a “scale” and that the scale change can be “reversed.” Let us first consider the case where the semi-ring $R$ is $\mathbb{R}$ or $\mathbb{C}$. In this case, $R$ forms a field and thus any non-zero element is a unit. It is easy to show that in this case strong equivalence implies abstract equivalence. To see that strong equivalence is a special case of abstract equivalence, let us consider mappings $f$ and $g$ to be defined at the origin “0” of the field (i.e. we ignore the translation between the mappings). If non-trivial mappings $f$ and $g$ are strongly equivalent, then $f = cg$, where $c$ is a non-zero real or complex number. We observe that $c$ is a unit and therefore its inverse $c^{-1}$ exists. Finally, we note that the strongly equivalent mappings $f$ and $g$ are abstractly equivalent.

We now extend the discussion to semi-ring $R$ given by $\mathbb{R}^n$ or $\mathbb{C}^n$. We note that $R$ forms a vector space over $\mathbb{R}$ or $\mathbb{C}$. We recall that the orthogonal linear operator is a necessary and sufficient condition for strong equivalence under the correlation function. Moreover, we note that the set of orthogonal linear operators forms an orthogonal group $O(n)$ given by $O(n) = \{ M \in \mathbb{C}^{n \times n} : M^H M = I \}$, where $I$ denotes the identity operator. The orthogonal group contains the special orthogonal group $SO(n)$ which represents usual rotations and is given by $SO(n) = \{ M \in \mathbb{C}^{n \times n} : M^H M = I \text{ and } \det(M) = 1 \}$. Finally, we observe that if $f$ and $g$ are strongly equivalent under the correlation function, then $f = Mg$, where $M \in O(n)$. Since $M$ is in the orthogonal group $O(n)$, we note that it is a unit (i.e. $M^{-1} = M^H \in O(n)$). Therefore, we once again conclude that strong equivalence implies abstract equivalence.
2.4 Applications and Examples in Genomic Signal Processing

As we discussed in previous sections, approach of mapping the symbolic sequence to $\mathbb{R}^n$ is a widely adopted method for symbolic signal process. Therefore the consistency problem for results using different mappings always arises. In this section, we will apply our theory to genomic signal processing.

We conduct experiments on Human gene AD169 sequence (GenBank accession no. X17403). We calculate the correlation function as in (Equation 2.6) using two different mappings. The first one maps the $\mathcal{A} = \{A, T, G, C\}$ to the standard basis of $\mathbb{R}^4$ correspondingly. Then we

![Graphs](image-url)  
(a) Correlation coefficients for strong equivalence  
(b) Percentage of points preserving local extremes for weak equivalence

Figure 2. Two consistency measurements for correlation results using two mapping methods change with the growth of sequence length $N$ for AD169 DNA sequence.
use another mapping strategy, which maps $A$ to $(-1,0,0,0)$, $T$ to $(1,0,0,0)$, $G$ to $(0,1,0,0)$ and $C$ to $(0,-1,0,0)$. These are two widely used mapping methods (10), (11). In Figure 2(a), we show the changing of correlation coefficient between the two correlation results with growth of DNA sequence length $N$ and in (b) we show how the percentage of the points having same local extremum property in two results grows with $N$. The second mapping is not obtained by rotation of the first mapping. As a result, all these two metrics have a decreasing trend with the grown of length $N$.

The example shows the same trends for two metrics between two mappings. The similarity between the two results become less and less, which finally may lead to an inconsistent analysis results due to the fact that two chosen mapping methods are not equivalent for the correlation function. Thus it does not make sense to make comparison between the analysis result for a given gene sequence under these two mapping methods.

In Figure 4, we show the consistency measurements between the power spectrum under the previous two mapping methods. Although the equivalent transform we analyzed before does not mainly focus on power spectrum, we can still find that the power spectrum results using these two different mappings have the trend to be inconsistent. Since the correlation and power spectrum are widely used and pervasive in statistic analysis, it suggests the consistency problem should not be neglected when comparing analysis results.

Research on statistical properties of coding and non-coding regions in nucleotide sequences is an important topic in genomic signal processing (21; 22). We shall also conduct experiments on coding and non-coding regions of Human gene NOC2L using the two mapping methods
Figure 3. Percentage of points preserving local extremes for Fourier transform using three different maps changes with growth of sequence length $N$ for human gene AD169 sequences.

introduced earlier. As shown in Figure 5, the consistency measure between the correlation functions decays as the length $N$ increases. As a result of our analysis, we note that the correlation results under the two mappings are inconsistent in the long run. Furthermore, any comparison between the analysis results obtained by relying on these correlation functions becomes increasingly unreliable. Another interesting result can be observed in Figure 5, where the decay rate of the non-coding region is faster than the coding region. This phenomenon could be attributed to the fact that the coding regions can be viewed as more random than the non-coding regions. Nevertheless, the main conclusion that we draw our attention to is that the consistency of the correlation between non-equivalent mappings decays as the sequence length increases for both coding and non-coding regions.
Figure 4. The correlation coefficient of the power spectrum changes with growth of sequence length \( N \) for Human gene AD169 sequences under the two mapping methods.

We calculate the Fourier transform as defined in (Equation 2.24) on Human gene AD169 sequence. The first mapping is chosen as before, which maps the \( \mathcal{A} = \{ A, \ T, \ G, \ C \} \) to the standard basis of \( \mathbb{R}^4 \) respectively. Then we use second mapping strategy, which maps \( A \) to \((0.9912, 0.1322, 0, 0)\), \( T \) to \((0.8367, -0.239, 0.1195, 0.4781)\), \( G \) to \((-0.7505, -0.5361, -0.2144, 0.3216)\) and \( C \) to \((0.7804, -0.5103, -0.2401, -0.2701)\). The third strategy maps \( A \) to \((\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0)\), \( T \) to \((0, \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})\), \( G \) to \((-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0)\) and \( C \) to \((0, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})\). We have normalized the mappings so that it will not change the energy of the result. The second mapping is not obtained by rotating the first mapping. While the third mapping is obtained by rotating the first mapping. In Figure 3, we show the weak equivalence metrics for Fourier analysis results using these three different mappings. The consistency between results using the first and second mapping becomes less and
Figure 5. Correlation coefficients for strong consistency measure of correlation functions changes with growth of $N$ for coding and non-coding regions of Human gene NOC2L.

less. While the results using the first and third mapping are completely consistent. In Figure 6, we show the analysis result of the three mappings. As we showed before, the results using the first and third mapping are exactly same, since rotation is a strongly equivalent transformation. We can find many differences between the results using the first and second mapping, especially at the peaks. We also calculate the correlation coefficient between them, which is 0.82. The peak here means the repeat pattern of some periodic sequences, however, since we have shown that the mapping is not equivalent here, it makes no reason to debate on possible conflicting analysis results for this gene sequence.

In all of the experiments conducted we observe that rotation serves as the unique equivalent transformation for the correlation function. Rotation also provides a strongly equivalent transformation for Fourier and spectrum analysis. Mappings which are not equivalent lead to
inconsistent results as the sequence length $N$ increases. However, we must point out that the opposite may not be true: specifically, for a fixed-length sequence, the consistency between mappings does not necessarily decay as the difference between mappings increases, measured in the sense of rotation equivalence, i.e. the similarity between the first mapping and any mapping obtained by rotating the second mapping.

2.5 Conclusion

In this chapter, we presented a novel framework for analysis of the equivalence of distinct numerical mappings of symbolic sequences undergoing a transformation by an operator. We introduced a strong equivalence property that demands perfect correlation between the transformations of distinct numerical representations. We also characterized the weak equivalence property which requires the preservation of the extrema in the transformation of the numerical
representations. We studied the mapping equivalence theory for general operators by using Taylor’s approximation. Moreover, we focused on first- and second-order operators such as the correlation function and Fourier transform. Furthermore, we derived the largest class of equivalent mappings which lead to consistent results when undergoing transformation by a class of operators. We demonstrated that rotation plays an important role in characterization of equivalence between distinct mappings. We subsequently derived a class of operators which is equivalent under rotations. We also introduced an abstract mapping model and extended the notion of equivalence to a more general algebraic structure. We presented simulations of the mathematical and statistical properties of genomic sequences in order to demonstrate the implications of the proposed mapping equivalence theory. Our results suggest that one of the reasons for inconsistency in the analysis of genomic data reported in the theoretical biology literature as well as many other related areas can be attributed to incompatibility of the numerical representation of symbolic data.
CHAPTER 3

DYNAMICS, STABILITY AND CONSISTENCY FOR SYMBOLIC SEQUENCES PROCESSING

3.1 Introduction

As we point out in previous chapter numerous mapping strategies have been proposed for different areas, indeed it is impossible to determine which mapping is preferable. Furthermore, it is conceivable that distinct mappings could lead to contradictory conclusions. There is always an issue of the consistency for the analysis results under different mappings. In previous chapter, we proposed the concept of mapping equivalence and showed the equivalent transform for certain classes of operators including correlation function and Fourier transform.

However, operators as de-noising filter, smoothing filter and certain algorithm may be applied many times in processing the data. The consistency question rises again naturally in this situation. Another important question in this situation would be whether the analysis result is robust and stable to the perturbations of the data. Due to the noises and errors in acquisition of the biological data sequences, the mapping equivalence or the stability issue under the iterations of operators can not be neglected.

In this chapter, we provide an approach for analyzing the consistency and stability for different mappings under iterations of operator. In Section 3.2, we first define the mapping equivalence and introduce the concept of Fatou and Julia Set. We also show the necessary and
sufficient condition for consistency under iteration of affine operator. In Section 3.3, we give the definition of the stability, which can be shown to be a special case of mapping equivalence. In particular, we establish the connection between stability between Fatou and Julia set. In Section 3.4, we present experimental results which illustrate the theoretical results on genomic sequence. Finally, we provide a brief summary and discussion of our results in Section 3.5.

3.2 Dynamics and Mapping Consistency Under Iterations

Given biological data sequence \( \{a_i\}_{i=0}^{n-1} \), where \( a_i \in \mathcal{A} \). The set \( \mathcal{A} \) could be collection of nucleotides, amino acids, etc. \( f \) is a mapping from \( \mathcal{A}^n \) to \( \mathbb{C}^N \), i.e. \( f : \{a_i\}_{i=0}^{n-1} \mapsto z, \ z \in \mathbb{C}^N \). In particular, if we have a mapping method \( \tilde{g} : \mathcal{A} \mapsto \mathbb{C}^k \), then it naturally induces the map \( g : \{a_i\}_{i=0}^{n-1} \mapsto z, \ z \in \mathbb{C}^{nk} \), where \( \left( \left[ z \right]_{jk+1}, \left[ z \right]_{jk+2}, ..., \left[ z \right]_{jk+k} \right)^T = g(a_j), \ j = 0, 1, ..., n-1 \). Therefore, for a given symbolic sequence and a mapping method \( f \), the corresponding numerical sequence is a point in \( \mathbb{C}^N \). We denote this point as \( z_f \). Let \( \Phi : \mathbb{C}^N \mapsto \mathbb{C}^N \) be a holomorphic (analytic) operator. In this thesis we will assume \( \Phi \) is polynomial, i.e. \( (\Phi(z))_i = P_i(z_1, z_2, ..., z_N), i = 1, ..., N \), where \( P_i \) is a polynomial. From Taylor’s theorem we know that any holomorphic map can be approximated by polynomials.

3.2.1 Mapping Consistency Under Iterations

Since we need the analysis result to be consistent in some sense, we first define different concepts of mapping equivalence.

Definition 5 Given a genomic sequence and two mapping \( f \) and \( g \), we say \( f \) and \( g \) are asymptotically equivalent if

\[
\lim_{n \to \infty} \| \Phi^{on}(z_f) - \Phi^{on}(z_g) \| = 0
\]  

(3.1)
\( f \) and \( g \) are called \( M \)-boundedly equivalent, if

\[
\sup_n \| \Phi^n(z_f) - \Phi^n(z_g) \| < M. \tag{3.2}
\]

\( f \) and \( g \) are called \( n \)-th equivalent, if

\[
\| \Phi^n(z_f) - \Phi^n(z_g) \| = 0 \tag{3.3}
\]

We say \( f \) and \( g \) are equivalent if they are \( k \)-th equivalent for any \( k \in \mathbb{N} \).

In study of dynamics of equivalence, the concept of Fatou and Julia set play a fundamental role. There are several different definitions of Fatou and Julia set \((29; 30)\). We will use the definition below. Before that, we first introduce the notion of normality.

**Definition 6 ((49))** A collection of holomorphic map \( \mathcal{F} \) is called normal if every infinite sequence of maps from \( \mathcal{F} \) either has a locally uniformly convergent subsequence or a subsequence diverges locally uniformly.

**Definition 7 ((29))** The domain of normality \( F \) of \( \mathcal{F} = \{ \Phi^n \} \) is called Fatou set. Its complement

\[
J = \mathbb{C}^N \setminus F \tag{3.4}
\]

is called Julia set.

We define the basin of infinity as set of all points which have norms go to infinity under iteration.
The connected components of Julia (Fatou) set are called Julia (Fatou) components.

We will see later the Julia set represents the chaotic behaved points and points in Fatou set show rational behavior. We can show the following propositions about the Fatou set. The proof of the following two propositions in one-dimensional case appears in (29).

**Proposition 7** A point \( z \) is in Fatou set if \( z \) is in the basin of infinity \( B \).

**Proof 10** It is straightforward to see if \( z \in B \), then there exists a neighborhood \( U_z \) of \( z \), such that

\[
\lim_{n \to \infty} \| \Phi^o_n(z) \| \to \infty, \forall z \in U_z
\]  

(3.5)

So we have \( z \in F \).

**Proposition 8** Fatou (Julia) component is invariant. i.e. the operator maps one component to another component.

**Proof 11** It follows from the fact \( \mathcal{F} = \{ \Phi^o \} \) and \( \mathcal{F} = \{ \Phi^{(n+1)} \} \) have the same domain of normality and \( \Phi \) is continuous.

For \( z_f \) and \( z_g \), if only one of them is in the basin of infinity, it is obviously that \( f \) and \( g \) are not boundedly or asymptotically equivalent. If both are in the basin of infinity, although theoretically we can examine the equivalence, however, from computational point of view, the point diverges very fast under polynomial iterations. After a few rounds of iterations, the numerical results will overflow. In this case, the equivalence or even analysis result turns out to be meaningless. Braverman and Yampolsky (50) showed the Julia set of certain types of
polynomial can not be computed by any Oracle Turing Machine. If one of the point is in Julia set, it may not able to figure out the equivalent mapping class of the given maps. We will classify all the mappings falling in all these situations as the \textit{computationally chaotic mapping class}. In general, from computational point of view, it is futile or meaningless to find the equivalent mappings of the element in computationally chaotic mapping class. Therefore, the only interesting case left would be if both points are in Fatou set.

\subsection*{3.2.2 Dynamics Under Iterations}

As for the simplest case if $\Phi$ is affine. We can show the following results for equivalent mapping.

\textbf{Theorem 5} If $\Phi(z) = Az + b$, then a pair of maps $f$ and $g$ is asymptotically equivalent for any genomic sequence if and only if the spectral radius $\rho(A) < 1$.

Any pair of maps is boundedly equivalent for any genomic sequence if and only if either $\rho(A) < 1$ or $\rho(A) = 1$ and all the eigenvalues have index $\leq 1$.

\textbf{Proof 12} If for any $x$ and $y$, $\|\Phi^{on}(x) - \Phi^{on}(y)\| \to 0$ as $n \to \infty$, then we have

$$\lim_{r \to \infty} A^r = 0 \quad (3.6)$$

Denote the $n \times n$ Jordan matrix with diagonal 0 as $J_n$. Consider the Jordan canonical form of $A$.

$$UAU^{-1} = \oplus_{i=1}^l \oplus_{j=1}^k (\lambda_i I_{m_{ij}} + J_{m_{ij}}) \quad (3.7)$$

where $m_{ij}$ is the size of $(i,j)$ jordan block.
Using the matrix function (44), we have that
\[
    A^r = U^{-1}(\bigoplus_{i=1}^k \bigoplus_{j=1}^k \sum_{p=0}^{m_{ij} - 1} \binom{r}{p} \lambda^{r-k} J^{p}_{m_{ij}}))U \tag{3.8}
\]

So we must have the necessary and sufficient condition that \( \rho(A) < 1 \).

If we have any two maps are boundedly equivalent, then we have
\[
    \|A^r\| < M \tag{3.9}
\]

for any \( r \).

Follow from (Equation 3.8), we have either \( \rho(A) < 1 \) or if for some  \(|\lambda_i| = 1\), we must have there is only one Jordan block for that \( \lambda_i \), i.e. \( \text{index}(A) = 0 \) if \( \rho(A) = 1 \).

### 3.2.3 Kobayashi Metric and Iteration

Though in applications, Euclidean metric is the most widely used metric. However, it is often easier to work with the Kobayashi metric (51) than the Euclidean metric for investigating the quasi-stability. For complex manifold \( M \), one can construct a differential metric \( F_M : T(M) \to \mathbb{R} \).

**Definition 1** \( F_M(\xi_x) := \inf \left( \frac{1}{r} : \exists f : D(r) \to M \text{ such that } f(0) = x \text{ and } df(\frac{\partial}{\partial x})_0 = \xi_x \right) \)

Once we have this differential metric, we can construct a pseudo metric called *Kobayashi pseudo metric*. 
Definition 2 \( d_K(x, y) := \inf_{\gamma} \{ \int_{a}^{b} F_M(\dot{\gamma}(t)) \, dt \} \) where \( \gamma : [a, b] \to M \) is a piecewise \( C^\infty \) curve connecting \( x \) and \( y \).

A complex manifold is called **hyperbolic** if the Kobayashi pseudo metric \( d_K \) is a metric. (51) is referred for the details of construction and properties of Kobayashi metric and hyperbolic manifold. One of the fundamental theorem for hyperbolic manifold would be the non-increasing principle (51).

**Theorem 6 (Non-increasing Principle)** If \( M \) is hyperbolic, then for any holomorphic function \( \Phi \) we have,

\[
d_K(\Phi(x), \Phi(y)) \leq d_K(x, y), \quad x, y \in M
\]

We can show the following theorems on certain class of operator. Noted in the projective space \( \mathbb{P}^n \), two point are identified if one is a constant multiple of the other, i.e, \( z_f z_g \) if \( z_f = \lambda z_g \), where \( \lambda \in \mathbb{C} - \{0\} \). It would be more convenient to work in \( \mathbb{P}^n \) for checking the consistency, since if two results are multiple of each other, it is conceivable that these two results contains same information. Recall the a polynomial map \( f \) is non-degenerate if \( f^{-1}(0) = 0 \). For a non-degenerate homogenous polynomial from \( \mathbb{C}^{n+1} \) to \( \mathbb{C}^{n+1} \), it natually rises to map from \( \mathbb{P}^n \) to \( \mathbb{P}^n \) through the projection map \( \pi : \mathbb{C}^{n+1} \to \mathbb{P}^n \). So from now on we will assume the underline space is \( \mathbb{P}^n \).

Ueda showed the following important result about the Fatou component for non-degenerate homogenous polynomial (52).
Theorem 7 If \( \Phi : \mathbb{P}^n \to \mathbb{P}^n \) is non-degenerate homogenous polynomial on \( \mathbb{P}^n \), then its Fatou components are hyperbolic.

We can have the following results.

Theorem 8 If \( \Phi \) is non-degenerate homogenous polynomial, \( z_f \) and \( z_g \) are in Fatou set, if \( d_K(z_f, z_g) < M \) then \( f \) and \( g \) will be \( M \)-boundedly equivalent under \( d_K \) metric. In particular if \( z_f \) and \( z_g \) are in the same Fatou component \( U \) where \( \Phi(U) = U \) and \( U \) is hyperbolic, then any two mappings \( z_f \) and \( z_g \) in this Fatou component is boundedly equivalent under Euclidean metric.

Proof 13 The first claim follows directly from the definition. Using the property the Kobayashi metric is continuous (51), the open subsets with respect to the \( d_K \) topology are also open in the Euclidean topology. For the converse, for a point \( x \), choose a relatively compact neighborhood \( U \) of \( x \), consider \( r = \min_{y \in \partial U} \{ d_K(x, y) \} \), notice \( r > 0 \) so the \( r \) ball in \( d_K \) topology is contained in \( U \). So we have the Kobayashi topology is equivalent to Euclidean topology. Together with the non-increasing principle we proved the second claim.

Theorem 9 If \( \Phi \) is non-degenerate homogenous polynomial, \( U \) is a Fatou component, and \( \Phi(U) = U \), if \( d_K(\Phi(x), \Phi(y)) < d_K(x, y) \) for any distinct \( x, y \in U \), then there exists a unique fixed point in \( U \) and any \( z_f \) and \( z_g \) in \( U \) are asymptotically equivalent.

Proof 14 Choose a point \( p_0 \), let \( p_n = \Phi^n(p_0) \). If \( \lim_{n \to -\infty} d_K(p_0, p_n) = \infty \), then \( \forall q_0 \in B(p_0, r) \), by triangle inequality, we have \( d_K(q_n, p_0) \geq d_K(p_n, p_0) - r \to \infty \). Get a contradiction.
So we must have \( \lim_{n \to \infty} d_K(p_0, p_n) < M \). Therefore \( \exists n(1) < n(2) < \cdots \), such that \( p_{n(j)} \to \hat{p} \). Let \( g_j = \Phi^{n(j+1) - n(j)} \) and \( r_j = d_K(\hat{p}, p_{n(j)}) \). We have

\[
d_K(g_j(\hat{p}), p_{n(j+1)}) \leq r_j \tag{3.11}
\]
\[
d_K(g_j(\hat{p}), \hat{p}) \leq r_j + r_{j+1} \tag{3.12}
\]

If \( M \) is hyperbolic and paracompact then \( \text{Hol}(M, M) \) is a normal family (51). Also By Stone’s theorem that every metric space is paracompact (53). We have \( \text{Hol}(U, U) \) is a normal family.

Let \( g \) be the accumulation point of \( \{g_j\} \). So we have

\[
g(\hat{p}) = \hat{p} \tag{3.13}
\]

We have

\[
f(\hat{p}) = f(g(\hat{p})) = g(f(\hat{p})) \tag{3.14}
\]

but \( g \) has only 1 fix point, so \( f(\hat{p}) = \hat{p} \).

It also follows easily from the strictly deceasing that the fix point is an attracting point.

3.3 Stability Under iterations

In previous sections, we introduce the different concepts of mapping equivalence and provide some results on the equivalent mappings under iteration. Another issue in the iteration is whether the result in stable to the perturbation of initial values. Due to the inevitable errors
and noises in acquisition of data, the stability issue cannot be ignored. The analysis result
would be much less compelling if it is obtained under some unstable mapping strategy. Usually
the biological data sequence would possess very long length, after the numerical mapping, a
few acquisition error would correspond to another numerical sequence near the true sequence.
Therefore the stability issue is equivalent to the question that whether small changes of the
given sequence will cause a small changes in the result.

**Definition 8** A mapping \( f \) is stable, if for any \( \delta > 0 \), there exists \( \epsilon > 0 \) such that for any point
\( z_g \) in the ball of radius \( \epsilon \), centered at \( z_f \) we have

\[
\| \Phi^{\circ n}(z_f) - \Phi^{\circ n}(z_g) \| < \delta, \forall n \in \mathbb{N}
\]  

(3.15)

In another word, all the mappings in the ball are \( \delta \)-boundedly equivalent.

We can show the following results about the stability.

**Theorem 10** If \( z \) is not in basin of infinity, \( z \) is stable if and only if \( z_f \) is in Fatou set.

**Proof 15** If \( f \) is stable, then \( \mathcal{F} = \{ \Phi^{\circ n} \} \) is eqi-continuous. Followed by Arzelà-Ascoli theorem,
we have \( z_f \in F \).

Conversely, we will show if \( f \) is not stable, we will have \( z_f \in J \). If it is not stable, we have
\( \epsilon_0 \) and \( n(1) < n(2) < \ldots \) such that for \( \{ \Phi^{\circ n(j)} \} \), \( \exists \{ z_j \} \rightarrow z_0 \), we have

\[
| \Phi^{\circ n(j)}(z_j) - \Phi^{\circ n(j)}(z_0) | > \epsilon_0
\]  

(3.16)
Assume \( z_f \in F \), we have a neighborhood \( U \) of \( z_f \), such that \( \{\Phi^{on(j)}\} \) has a uniformly convergent subsequence. Let \( h \) denote the limit. Taking limit for (Equation 3.16). We have

\[
|h(z_j) - h(z_0)| > \epsilon_0
\] (3.17)

for \( j \) large enough. But this contradicts to \( g \) is continuous. So \( z_f \in J \).

From theorem 10, we can see that Fatou set represents the good-behaved mappings. Any mapping close enough will be a boundedly equivalent mapping. On the contrary, for the mapping in the Julia set, no matter how close the mapping is, it may not even be a boundedly equivalent mapping.

3.4 Genomic-Sequence Analysis

We conduct the experiments on human gene AD169 sequences (GenBank accession no. X17403) and rhodopsin gene sequence (GenBank accession no. U49742). We consider the operator \( \Phi \) as a non-linear smoothing filter defined as follow,

\[
\Phi(z_1, z_2, \ldots, z_N) = \left( \frac{z_1^2 + z_2^2}{2}, \ldots, \frac{z_i^2 + z_{i+1}^2}{2}, \ldots, \frac{z_N^2}{2} \right)
\] (3.18)
We consider the mapping \( \tilde{f} \) as in (8),

\[
\tilde{f}(a) = \begin{cases} 
1 & \text{if } a = A \\
-1 & \text{if } a = T \\
i & \text{if } a = G \\
-i & \text{if } a = C 
\end{cases}
\]  

(3.19)

This is one of the widely used mappings. We denote the induced mapping point as \( z_f \).

In Figure 7, we show the slices of Julia and Fatou set of \( \Phi \) at \((z,1,1,...,1), (z,i,i,...,i)\) and \((z,0.25 + 0.75i,...,0.25 + 0.75i)\). Julia set commonly possesses a fractal shape and could be connected or disconnected.

It can be shown the Fatou component \( U \) containing origin satisfies all the assumptions in theorem 9. Therefore any two mappings in \( U \) will be asymptotically equivalent and \( z_f \in U \) for both human gene AD169 sequence and rhodopsin gene sequence. We consider the following perturbed mapping \( \tilde{f}' \),

\[
\tilde{f}' = \tilde{f} + \Delta z
\]

(3.20)

where \( \Delta z \in \mathbb{C} \). In Figure 8, we show the slice of the Fatou component \( U \) with \( z_f \) at origin and \( \Delta z \) is varying in the ball of radius 0.1, centered at 0. The white area is in the Fatou component.

In Figure 9, we show how Euclidean distance for two arbitrarily chosen mappings which are in the previous Fatou component \( U \) changes with the number of iterations for human gene AD169 sequence. As we can see the distance converges to 0 with the increase of number of iterations.
In Figure 10, we show the case how Euclidean distance for two mappings changes with the number of iterations. One is in the previous Fatou component $U$ and the other is not. As we can see the distance diverges with the increase of number of iterations.

### 3.5 Conclusion

In this chapter, we provide a method for analyzing the consistency between different mappings under iterations of operator. We define different mapping equivalence concepts including
asymptotical, bounded and n-th iteration equivalence. We provide the necessary and sufficient condition for consistency under iteration of affine operator. We present a few results on the equivalent mappings based on the concept of Fatou and Julia Set. We give the definition of stability under iteration of operator and show the stability issue can be viewed as a special case of mapping equivalence. We also establish the connection of stability to Fatou and Julia set. Finally, we conduct experiment on human gene AD169 sequence and rhodopsin gene sequence under a popular mapping method. We show that it is stable for a smoothing filter and illustrate its equivalent mapping class. In the future, we will study the dynamics and consistency problem where there is a composition of different operators.
Figure 9. Consistency case: The illustration of how Euclidean distance for two mappings which are in the previous Fatou component $U$ changes with the number of iterations for human gene AD169 sequence.

Figure 10. Inconsistency case: The illustration of how Euclidean distance for two mappings, for which one is in the previous Fatou component $U$ and the other is not, changes with the number of iterations for human gene AD169 sequence.
CHAPTER 4

STOCHASTIC CONTROL FOR REGULATORY NETWORK

4.1 Introduction

The genetic regulatory modeling is aimed to describe the dynamics of gene networks, which is useful to identify potential drug targets or alter the system evolutions in a desired manner. It is well known that many diseases such as cancer and tumor are due to the facts of cells improper proliferation or misfunction of certain genes.

There are many models for the gene regulatory networks based on ordinary differential equations, partial differential equations, Boolean network, stochastic equations, etc (54). The probabilistic Boolean network (PBN) (32) is one commonly used model for genetic regulatory modeling. The purpose of intervention of the PBN is to reduce the probability to visit the unwanted states. Various methods in regulatory network intervention have been proposed. Shmulevich, et al proposed a one-time intervention strategy based on first-passage time (55). Datta et al proposed a method based on Markovian control (56). Methods using the Markov decision process have also been proposed (57; 34). One major difference between the one-time intervention and Markovian intervention is that the Markovian intervention is based on certain average concept. While the short-time intervention may not be very effective in the long run of the dynamics.

60
The PBN can be modeled as a Markov chain model under certain assumptions (33). The perturbed gene can be seen as the external control of the Markov process, which can be modelled as a Markov decision process (MDP). The optimal control policy can be solved by dynamic programming (35). In these approaches of intervention in PBN, the perturbed the genes are modeled as the external control and the optimal control policy is solved by maximizing the reward function. However, in many situations, multiple controls may be desired to obtain various optimalities for different purposes. For example, in drug or gene therapy, the patient may need to be treated for various diseases simultaneously based on vary large regulatory network. While the interactions in the network are difficult or intractable to understand. In these situations, a single control is not adequate to describe the interventions. In multiple controls situation, the independent and blind maximization of reward function for each control will not achieve the original aim since the dynamics of the system are bound together. Instead of solving a single optimization problem as in MDP, proper concept of solution in these problems i.e. the Nash equilibria is required to be defined. As we will see later in the paper, the classical MDP falls into a special case of the non-cooperative stochastic game (NCSG).

In this chapter, we propose the NCSG model for control of the genetic regulatory networks and formulate the intervention problem into solving the Nash equilibrium. In section 4.2. we first provide the necessary background of the PBN and introduce the NCSG model. In section 4.3, the MDP is shown to be a special case of NCSG and the solving methods are provided. In section 4.4, the definition of Nash equilibrium has been proposed and the existences for both infinite and finite horizon cases have been proven. In section 4.6, we provide the numerical
example for our NCSG model on the mammalian cell cycle network. We also compare the results under the Nash equilibrium and independent MDP (IMDP) method. Finally, we provide a brief summary and discussion of our results in Section 4.7.

4.2 The Models for Regulatory Network

The PBN (32) consists of $n$ nodes $N = \{x_i\}_{i=1}^n$, where $x_i \in \{0, 1\}$, a sequence of vector valued predictor functions $\{f_j\}_{j=1}^r$. The $x_i$ represents the expression level of gene with “0” meaning OFF and “1” meaning ON. The predictor functions $f_j = (f_{j1}, \ldots, f_{jn})$ is a predictor of genes 1 through $n$, when the network $j$ is selected. Note that the Boolean network is deterministic if the predictor functions are given. We have the switching probability $q$ to describe the randomness of the network. If the switch is on, the predictor functions are chosen randomly according to probability measure $\{p_l\}_{l=1}^q$. The vector $x(t) = (x_1(t), \ldots, x_n(t))$ can be seen as the binary expansion of $\{0, \ldots, 2^n - 1\}$, which is called Gene-activity profile (GAP).

For the intervention, we assume there are $K$ controls. In the context of Game theory, we refer them to be $K$ players. For each player $k$, it controls genes $a_k = \{g_{k1}, \ldots, g_{km_k}\}$. Therefore the action set for player $k$ is $A_k = \{0, 1, \ldots, 2^{m_k} - 1\}$. We denote $a = \prod_{k=1}^K a_k$ and $A = \prod_{k=1}^K A_k$. We also assume the genes each player controls are disjoint. At every epoch $t$ the network updates and every player make decision at the same time, i.e. each player only knows his own decision.
Let \( s_t = (x_1(t), \ldots, x_n(t)) \) denote the network state. The system can be modelled as the controlled Markov process. The transition probability is defined as,

\[
P\{s'\mid s, a\} := P\{s_{t+1} = s'\mid s_t = s, a_1 = a^{(1)}, \ldots, a_k = a^{(k)}\} \tag{4.1}
\]

If the initial state \( s_0 \) is given, then it fixed a stochastic process. Let \( r_k^e(s_t, a_1, \ldots, a_k) \) be the immediate reward function for player \( k \) at epoch \( t \). So we have the expected reward function \( R_k^e = E_{s_0}(r_k^e) \). In this thesis, we mainly focus on the finite and infinite horizon discount average reward functions:

\[
V_k^e(L) := \sum_{n=1}^{L} \beta^{n-1} R_n^k \tag{4.2}
\]

\[
V_k^e := \sum_{n=1}^{\infty} \beta^{n-1} R_n^k \tag{4.3}
\]

The policy \( \pi_k = \{\pi_{k1}, \pi_{k2}, \ldots\} \) for player \( k \) is a sequence of probability distributions such that \( \pi_{ki} \) is the decision distribution on the action set \( a_k \). A policy is pure, if all the decisions in a policy are deterministic. If \( \pi_{ki} \) is independent of time \( i \), then the policy is stationary.

4.3 The Single Player Case

For a single controller intervention, we can model it as a Markov decision process. The external control is applied on genes \( a = \{g_1, \ldots, g_k\} \). Therefore the action set is \( A_k = \{0, 1, \ldots, 2^k - 1\} \).
Let \( s_t = (x_1(t), \ldots, x_n(t)) \) denote the network state. The system can be modelled as a Markov process with transition probability controlled by the action of controller. The transition probability is defined as,

\[
P\{s_{t+1}|s_t, a\} := P\{s_{t+1} = s'|s_t = s, a_1 = a^{(1)}\}
\] (4.4)

The MDP is a special case of the Markov chain with transition probability controlled by the controller. If the initial state \( s_0 \) is given, then it fixed a stochastic process. For each stage of the process, we will have a reward function to provide the numerical description of current state. For example, the state which could lead to malicious status will have low reward value. Let

\[
r_t(s_t, a_1) : S \times A_k \to [0, \infty)
\] (4.5)

be the immediate reward function at epoch \( t \) for state \( s_t \). For the controller, the action it chooses at each time step \( t \) will be denoted by

\[
f_t : S_t \to A_t.
\] (4.6)

More generally, we can have a probability measure on \( f_t \), which make the action strategy random. With abuse of notation, we still denote \( f_t : S \times A_t \to [0, 1] \) as the strategy. Note that \( f_t \) is a probability measure on \( A_k \), i.e. \( \sum_a f_t(s_t, a) = 1 \). We call

\[
f = \{f_1, f_2, \ldots, \}
\] (4.7)
the policy. A strategy is called pure if \( f(s, \cdot) \) is deterministic for all \( s \in S \). A policy is called stationary if \( f_1 = f_2 = \cdots \).

Once we fix a strategy at time \( t \), we can see that the transition probability is determined as

\[
P\{s_{t+1}|s_t, f_t\} := \sum_{a \in A_t} P\{s_{t+1} = s'|s_t = s, a_1 = a\} f(s_t, a) \tag{4.8}
\]

After we introduce the random policy, we can calculate the expected reward function \( R^k_t = E_{s_0}(r^k_t) \) give the initial state \( s_0 \). The general total reward will a form of average of \( R^k_t \). There are two kinds of rewards, discount and limiting average (58). In this paper, we mainly focus on the discount model.

In the discount total reward case, we will introduce a discount factor \( \beta \in (0, 1) \). Each expected reward function at time \( t \) will be shrink by factor \( \beta^t \). We have the finite and infinite horizon discount average reward functions:

\[
V^L := \sum_{n=1}^L \beta^{n-1} R^k_n \tag{4.9}
\]

\[
V := \sum_{n=1}^{\infty} \beta^{n-1} R^k_n \tag{4.10}
\]
Therefore the single control intervention can be formulated as the following problem:

$$\max V(f)$$

Subject to: $f \in \mathcal{F}$ \hspace{1cm} (4.11)

where $\mathcal{F}$ is the solution class such as all stationary policies, etc.

Problem Equation 4.11 is easy to put into the dynamic programming problem, namely if we know the optimal policy up to time $L - 1$ then we can simply maximize the expected reward for the remaining one stage.

In the infinite horizon case, the problem can be solved by the following linear programming (58).

$$\max \sum_s \sum_{a=0}^{2^m-1} r(s, a) z(s, a)$$

subject to:

$$\sum_s \sum_{a=1}^{2^m-1} \{\delta(s, s') - \beta \mathbb{P}\{s'|s, a\}\} z(s, a) = \frac{1}{\pi}$$

$$z(s, a) \geq 0$$

(4.12)

The optimal stationary policy is given by

$$\pi = \frac{z^*}{\sum_a z^*}$$

(4.13)

where $z^*$ is the solution of the linear programming (Equation 4.12).
4.4 The Non-cooperative Multiple Players Case

The NCSG is the generalization of MDP. In multi-player case, although the goal for each player is to maximize his average reward function, however, the independent IMDP solution for each player is not adequate, i.e. each control works as if it were the only control and treats the problem as the MDP. Because the reward function and transition probability are coupled together by the decisions of all players.

Unlike the case in MDP, where one reward function is to be maximized. In NCSG, all players are interested in maximizing their individual average reward functions. Therefore a proper solution concept is needed to be defined. The concept of Nash equilibrium in static game can be extended as the equilibrium in NCSG.

**Definition 9** A policy $\pi^* = (\pi_1, \ldots, \pi_K)$ is the infinite horizon discount Nash equilibrium (NE) point if the following inequality holds,

$$V_k(\pi^*) \geq V_k(\mu_k, \pi^*_{-k}) \quad (4.14)$$

for any $\mu_k \in \pi_k$ and any $k \in K$. The $\pi_{-k}$ denotes the policies of all players except player $k$.

Similarly, we also have the definition for finite horizon discount Nash equilibrium point.

**Definition 10** A policy $\pi^* = (\pi_1, \ldots, \pi_K)$ is the $L$-step finite horizon discount Nash equilibrium (NE) point if the following inequality holds,

$$V_k(L)(\pi^*) \geq V_k(L)(\mu_k, \pi^*_{-k}) \quad (4.15)$$
for any $\mu_k \in \pi_k$ and any $k \in K$. The $\pi_{-k}$ denotes the policies of all players except player $k$.

The existence of infinite horizon discount equilibrium point has been proven by A. Fink (59). By using the Nash theorem (60), we can also show the existence in finite horizon case.

**Lemma 9** Any policy $\pi_k$ is a convex combination of the pure policies in finite horizon discount game.

**Proof 16** We show the scheme to construct such convex combinations for $L = 2$. First consider the $\pi_{k1}$, it is straightforward to see

$$\pi_{k1} = \sum_{i=1}^{M} c_i^1 \mu_{k1}^i$$ (4.16)

where $\sum_i c_i^1 = 1$. $\mu_{k1}^i$ denotes the pure policy for player $k$ at epoch 1, which has probability 1 on $i$-th gene. Therefore consider the policy

$$\pi^{(j)} = \sum_{i=1}^{M} c_i^1 (\mu_{k1}^i, \mu_{k1}^j)$$ (4.17)

The first action in this policy $\pi^{(j)}$ is the same as $\pi_{k1}$, the second action is the pure policy which has probability 1 on the $j$-th gene. So

$$(\pi_{k1}, \pi_{k2}) = \sum_{j=1}^{M} c_j^2 \pi^{(j)}$$ (4.18)

where $\sum_i c_i^2 = 1$. Therefore

$$(\pi_{k1}, \pi_{k2}) = \sum_{i,j=1}^{M} c_i^2 c_j c_i^1 (\mu_{k1}^i, \mu_{k1}^j)$$ (4.19)
We see that the policy is convex combination of pure policies. It follows by mathematical induction for arbitrary stage $L$.

**Theorem 11** The finite horizon discount game has a Nash equilibrium point.

**Proof 17** For any given stage $L$ and initial state $s_0$, since the states and actions sets are finite, there are finitely many pure policies. By lemma 9, any policies is convex combination of pure policies. Therefore the set of all policies is a closed, bounded, and convex polyhedron. By Nash theorem (60), the equilibrium exists for finite horizon discount game.

4.5 **Solving Nash Equilibrium**

Under certain additional conditions, finding Nash equilibrium could be solved by linear programming or non-linear programming method. For example, if we have a two-player zero-sum game, i.e. $K = 2$ and $r^1 = -r^2$. This game can be interpreted as the matrix game and can be solved by Newton’s method (58).

In general, solving NE has been shown to be Polynomial Parity Arguments on Directed graphs (PPAD) complete (61). Therefore we do not have efficient algorithm to solve NE. Therefore we will use the concept of approximate NE and propose a constructive method for solving the approximate NE.

4.5.1 **Approximate Nash Equilibrium**

**Definition 3** A policy $\pi^*$ is called $\epsilon$-NE, if we have

$$V_k(\pi^*) \geq V_k(\mu_k, \pi_{-k}^*) - \epsilon$$ (4.20)
for any $\mu_k \in \pi_k$ and any $k \in K$.

If we normalize all the cost within $[0,1]$, we can construct a scheme which achieve at least $\frac{1}{2}$-approximation.

**Theorem 12** For 2-person game, if there exists $(i,j,k)$ such that for the cost $r^k(s,a_1,a_2)$, we have $j = \arg \max_{j'} r^j_2(s,i,j')$ and $k = \arg \max_{k'} r^k_1(s,k',j)$. Then the policy

$$x^* = \left(\frac{1}{2}\delta_i + \frac{1}{2}\delta_k, \ldots \right)$$

$$y^* = (\delta_j, \ldots)$$

(4.21)

will be at least $\frac{1}{2}$-approximate NE.

**Proof 18** If the assumption is satisfied, then for the constructive policy (Equation 4.21) we have,

$$\sum_s (\delta_i r^1(s)y^* - x^* r^1(s)y^*) f(s) < \epsilon$$

(4.22)

$$\sum_s (r^1(s,l,j) - \frac{1}{2} r^1(s,i,j) - \frac{1}{2} r^1(s,k,j)) f(s) \leq$$

(4.23)

$$\sum_s (\frac{1}{2} r^1(s,k,j) - \frac{1}{2} r^1(s,i,j)) f(s) \leq \frac{1}{2} r^1(s,k,j) \leq \frac{1}{2}$$

(4.24)

$$\sum_s (x^* r^1(s)\delta_i - x^* r^1(s)y^*) f(s) < \epsilon$$

(4.25)

$$\sum_s (\frac{1}{2} r^2(s,i,l) + \frac{1}{2} r^2(s,k,l) - \frac{1}{2} r^2(s,i,j) - \frac{1}{2} r^2(s,k,j)) \leq$$

(4.26)

$$\sum_s (\frac{1}{2} r^2(s,k,l) - \frac{1}{2} r^2(s,k,j)) f(s) \leq \frac{1}{2} r^2(s,k,l) \leq \frac{1}{2}$$

(4.27)
The constructive policy is only for 2-person game, however we have the following theorem for multiple-person game.

**Theorem 10** If for \((k-1)\) player we have an \(e\)-approximation, then there exists \(\frac{1}{2-\epsilon}\) approximation for \(k\) players.

**Proof 19** For the newly added player, he uses an arbitrary mixed policy with support on two action point with probability \(\alpha\). So for original players we have the approximation

\[
(1 - \alpha)e + \alpha \tag{4.28}
\]

For the newly added player we have the approximation

\[
1 - \alpha \tag{4.29}
\]

Equates two approximation, we have

\[
\alpha = \frac{1 - \epsilon}{2 - \epsilon} \tag{4.30}
\]

Plug (Equation 4.30) back to (Equation 4.29), we have the claim in the theorem.

Combine the two previous theorems, we have the following corollary.

**Corollary 11** For 3-player game, we have \(\frac{2}{3}\) approximation NE.
The strategy in theorem 10 has poorer and poorer performance with the increase of number of players.

**Proposition 12** Fix an \( \epsilon \)-approximation for \( n \) players, if we use the strategy in the theorem 10, then we have

\[
\lim_{n \to \infty} \epsilon_n = 1 \quad (4.31)
\]

**Proof 20** We have

\[
\frac{1}{2 - \epsilon} - \epsilon = \frac{(1 - \epsilon)^2}{2 - \epsilon} > 0 \quad (4.32)
\]

for all \( \epsilon \in [0, 1] \). So \( \{\epsilon_n\} \) is an increasing sequence. Therefore the limit \( a = \lim_{n \to \infty} \epsilon_n \) exists.

Taking limit of both side of the equation

\[
\epsilon_{n+1} = \frac{1}{2 - \epsilon_n} \quad (4.33)
\]

We have

\[
a = \frac{1}{2 - a} \quad (4.34)
\]

Solving the equation, we have \( a = 1 \), i.e.

\[
\lim_{n \to \infty} \epsilon_n = 1 \quad (4.35)
\]
4.6 Experiments on Mammalian Network

In this section, we conduct an experiment based on the mammalian cell cycle with a mutated phenotype. The cycle regulation is proposed in (62). We order these genes as $s = \{\text{CycD, Rb, E2F, CycE, CycA, Cdc20, Cdh1, Ubc, CycB}\}$. The state sets $\{s\}$ can be interpreted as the binary expansion of $\{0, 1, \ldots, 511\}$. Depending on the value of input CycD, we have two constituent Boolean networks of the PBN. We assume the two constituent networks have the same probability and the probability of switching is 0.01. In Table Table I, we show the logic relation of all these genes.

We have the Rb and CycA genes as the two controls $a_1$ and $a_2$ of the PBN respectively. We also assume the following reward functions.

$$r^1 = \begin{cases} 
10 & \text{if } a_1 = 0 \text{ and } (\text{CycD, Rb}) \neq (0,0) \\
0 & \text{if } a_1 = 1 \text{ and } (\text{CycD, Rb}) = (0,0) \\
2 & \text{otherwise}
\end{cases} \quad (4.36)$$

$$r^2 = \begin{cases} 
9 & \text{if } a_2 = 0 \text{ and } (\text{CycD, CycA}) \neq (0,0) \\
1 & \text{if } a_2 = 1 \text{ and } (\text{CycD, CycA}) = (0,0) \\
3 & \text{otherwise}
\end{cases} \quad (4.37)$$

In Figure 11. we show the average rewards under two scenarios. The first one is the PBN with only one control on gene Rb by using the linear programming method in section 4.3. The second scenario is the PBN has two control, but each control employs IMDP method. We can
<table>
<thead>
<tr>
<th>Gene</th>
<th>Predictors</th>
</tr>
</thead>
<tbody>
<tr>
<td>CycD</td>
<td>Input</td>
</tr>
<tr>
<td></td>
<td>(CycD ∨ CycE ∨ CycA ∨ CycB)</td>
</tr>
<tr>
<td>Rb</td>
<td>(Rb ∨ CycE ∨ CycA ∨ CycB)</td>
</tr>
<tr>
<td>E2F</td>
<td>(E2F ∨ Rb)</td>
</tr>
<tr>
<td>CycE</td>
<td>(E2F ∨ Rb)</td>
</tr>
<tr>
<td>CycA</td>
<td>(E2F ∨ Rb ∨ Cdc20 ∨ (Cdh1 ∨ Ubc)) ∨</td>
</tr>
<tr>
<td></td>
<td>(CycA ∨ Rb ∨ Cdc20 ∨ (Cdh1 ∨ Ubc))</td>
</tr>
<tr>
<td>Cdc20</td>
<td>CycB</td>
</tr>
<tr>
<td>Cdh1</td>
<td>((CycA ∨ CycB) ∨ Cdc20)</td>
</tr>
<tr>
<td>Ubc</td>
<td>(Cdh1) ∨ (Cdh1 ∨ Ubc ∨ Cdc20 ∨ CycA ∨ CycB)</td>
</tr>
<tr>
<td>CycB</td>
<td>(Cdc20 ∨ Cdh1)</td>
</tr>
</tbody>
</table>

**TABLE I**

**BOOLEAN FUNCTIONS OF MAMMALIAN CELL CYCLE.**

find in the illustration that the average reward under the second scenario is lower than the first one. The reason is that under the second scenario, the transition probability is coupled with the two controls together. In general the IMDP method with negligence of other controls will not achieve the optimality. The Figure 12. shows the same situation for control on gene CycA.

In Figure 13. and Figure 14, we compare the average rewards under the IMDP and the equilibrium. We can find in the illustrations, the rewards under the equilibrium are generally better than IMDP method case. Since the effects of other controls have been taken into account and no player can profitably deviate by any of his own actions, which serves as a solution concept under multiple players situation.
Figure 11. Illustrations of the average reward of the optimal intervention for control $a_1$ on gene Rb under the case only one control and the average reward of the optimal intervention for control on Rb using IMDP method with the other control $a_2$ on CycA ignored.

4.7 Conclusion

In this chapter, we propose the NCSG model for intervention of the genetic regulatory networks which extends previous MDP model into multiple controls case. We formulate the intervention problem into solving the Nash equilibrium. The definition of equilibrium has been proposed and the existences for both infinite and finite horizon cases have been provided. We also provide the numerical example for using our proposed NCSG model on the mammalian cell cycle network.
Figure 12. Illustrations of the average reward of the optimal intervention for control $a_2$ on gene CycA under the case only one control and the average reward of the optimal intervention for control on CycA using IMDP method with the other control $a_1$ on Rb ignored.

Figure 13. Illustrations of the average reward of the optimal intervention for control $a_1$ on gene Rb under the two controls case using IMDP method and the average reward of the optimal intervention for control on CycA under equilibrium
Figure 14. Illustrations of the average reward of the optimal intervention for control $a_2$ on CycA under the two controls case using IMDP method and the average reward of the optimal intervention for control on Rb under equilibrium.
CHAPTER 5

CONCLUSIONS AND FUTURE WORK

5.1 Summary and Contributions

In this thesis, we propose the concept of mapping equivalence theory for the numerical representation of symbolic data. We propose a framework for the analysis of different numerical mappings undergoing transformation by an analytic operator using Taylor’s expansion. Moreover, we emphasize the investigation of first- and second-order operators including the correlation function and Fourier transform. We show results under different equivalence concepts.

We extend the mapping equivalence theory for iterated operator. We provide a method for analyzing the consistency between different mappings under iterations of operator. We define different concepts of mapping equivalence. We show the necessary and sufficient condition for consistency under iteration of affine operator. We present a few theoretical results for the equivalent mappings on the concept of Fatou and Julia Set.

We propose NCSG model for control of the genetic regulatory networks and formulate the intervention problem into solving the NE. We show that the MDP is a special case of NCSG and the solving methods are provided. The definition of NE in this context has been proposed and the existences for both infinite and finite horizon cases have been proven. We provide a constructive method for solving the approximate NE.
5.2 Recommendations for Future Work

5.2.1 Consistency Criterion for Composition of Operators

In Chapter 1 and 2, we present the framework and method to analyze the mapping consistency for an operator. The strong and weak equivalent transform have been investigated for the correlation and Fourier transform. However, as one can see, the proof highly depends on the specific operator. In real application, it is common to have composition of different operators. Therefore we have

\[ \Phi = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_n \]  

(5.1)

In general, we will require a consistency concept for \( \Phi \). Therefore it is interesting to find a criterion for each operator \( \Phi_n \) such that if we find the consistent mappings for each \( \Phi_n \), we can conclude the \( \Phi \) will have certain class of consistent mappings. Indeed in Chapter 2, what we investigate is the case where \( n \to \infty \) and \( \Phi_1 = \Phi_2 = \cdots \). One can see in this case, some class of mappings which is not consistent under a single operator becomes consistent due to the dynamics involving.

5.2.2 Correlated Equilibrium and New Solution Concept

The NE is a solution concept which assumes independent behavior for each player. Albeit a successful concept in many aspects, it is in general very difficult to find a NE. As we see in Chapter 4, one remedy is to find the approximate-NE.
Another solution concept is the correlated equilibrium (CE) (63). Instead of assuming the independence among all the players, we just assume they have a joint distribution. Therefore the payoff functions become a system of linear equations. The CE will be intersection of these hyperplanes, which is convex by nature. The NE will be a special case of CE, actually the vertices of the polytope. The NE will be the case we have product of distributions for each individual player, i.e.

$$P_{n_1, n_2, \ldots, n_N} = \prod_i P_{n_i}$$

(5.2)

For CE, one can solve easily by linear programming. Therefore NE and CE are both extreme cases. CE is very easy to solve but it assumes the dependence of the players. NE assumes independence but is very difficult to solve. In real situation, the players are correlated in certain way between CE and NE. It is interesting to incorporate the true correlation information to have a proper solution concept. One can represent the correlation relation in a tree. The root of the tree will correspond to CE and the leaves correspond to NE. The true situation corresponds to the intermediate vertices. We propose another proper solution concept to capture this situation called Factor Graph-based Structural Equilibria (64).

5.2.3 Switching Control Game and Super-modularity

For NCSG, one can only resort to non-linear programming for solving the NE in general. For a handful of cases, we can solve the NE efficiently. One case would be the zero-sum game as we mentioned in Section 4.5. Another interesting situation will be the switching control
game. If we can partition the state space $S$ to disjoint subsets $S_1,\ldots,S_n$ such that in each $S_i$, the transition is controlled by individual player, i.e.

$$P(s'|s, a_1, \ldots, a_n) = P(s'|s, a_1), \forall s \in S_1$$

$$\vdots$$

$$P(s'|s, a_1, \ldots, a_n) = P(s'|s, a_n), \forall s \in S_n$$ \hspace{1cm} (5.3)

In this case, the NCSG can be decoupled to several MDP problem, which can be solved by linear programming.

For solving the MDP, although one can use the linear programming, however, in the case we can assume the super-modularity structure (65), the solution will possess a monotone structure which is a mix of two pure policies. In this situation, it will be even easier to find the solution than using linear programming.
CITED LITERATURE


VITA

NAME: Liming Wang

EDUCATION: University of Illinois at Chicago, Chicago, Illinois, USA
Department of Electrical and Computer Engineering

Ph.D., August 2011 (exp.)
- Advisor: Prof. Dan Schonfeld

University of Illinois at Chicago, Chicago, Illinois, USA
Department of Mathematics, Statistics, and Computer Science

M.S., Pure Math, August 2011 (exp.)
- Advisor: Prof. Stefan Wenger

Huazhong University of Science and Technology, Wuhan, Hubei, China
Department of Electronics and Information Engineering

B.S., Electrical Engineering, June, 2006
- Advisor: Prof. Hanqiang Cao

ACADEMIC EMPLOYMENT:

Postdoctoral Researcher February, 2011 - Present
Department of Electrical Engineering
Columbia University, New York, New York, USA

Research Assistant January, 2008 - December, 2010
Multimedia Communications Laboratory
University of Illinois at Chicago, Chicago, Illinois, USA

Research Assistant October, 2006 - September, 2007
Machine Vision Laboratory
University of Illinois at Chicago, Chicago, Illinois, USA

Research Assistant September, 2005 - June, 2006
Research Center of Wireless Broadband and Multimedia System
Huazhong University of Science and Technology, Wuhan, Hubei, China
Research Assistant February 2005 - August 2005
Motorola (Freescale) MCU and DSP Laboratory
Huazhong University of Science and Technology, Wuhan, Hubei, China

TEACHING EXPERIENCE:
Teaching Assistant September, 2006 - December, 2010
Department of Electrical and Computer Engineering
University of Illinois at Chicago, Chicago, Illinois, USA
Responsible for grading homework, instructing lab session, designing and supervising student projects, holding office hour, problem session and occasional lecture.

- ECE 267 Computer Organizations I, Fall 2006, Spring 2007, Fall 2008, Spring 2009
- ECE 340 Electronics I, Fall 2007; Spring 2008
- ECE 341 Probability and Random Processes, Fall 2007
- ECE 367 Microprocessor-Based Design, Spring 2010
- ECE 431 Analog Communication Circuits, Fall 2009; Fall 2010

HONORS:
- Student Presenter Award, Graduate College, University of Illinois at Chicago, 2009; 2010; 2011
- GSC Travel Award, Graduate Student Council, University of Illinois at Chicago, 2009; 2010
- University Best Thesis Award, Huazhong University of Science and Technology, 2006
- Outstanding Student Fellowship, Huazhong University of Science and Technology, 2002 - 2006
- Outstanding Student Scholarship, Department of Electronics and Information Engineering, Huazhong University of Science and Technology, 2002

PUBLICATIONS:

Journal Paper:

1. Optimal Quantization for Random Signal under $L_p$ norm. (with N. Piotto, etc), in preparation.


Conference Paper:


Thesis:


Transform Domain Based Super-resolution Reconstruction Algorithm from Compressed Video Sequence, Bachelor thesis, Huazhong University of Science and Technology, 2006

PROFESSIONAL ACTIVITIES:

Student Member, IEEE

Reviewer:
- IEEE Transactions on Information Theory
- EURASIP Journal on Advances in Signal Processing
- Journal of the Franklin Institute
- Various Conferences (ICASSP, SPAWC, etc)

TECHNICAL SKILLS:

- Database: Oracle, SQL Server, Access
- Programming Languages: C/C++, Matlab, Verilog HDL, Assembly (Intel/Motorola), HTML, CSS
- Software Packages: MATLAB, Maple, Visual C++, Protel, Orcad, Altera Quartus, Freescale CodeWarrior, \LaTeX{}, Microsoft Office, Dreamweaver