


Spring 2017

Dynamics of Gene Networks in Cancer Research

Paul Scott

Follow this and additional works at: <https://digitalcommons.georgiasouthern.edu/etd>

 Part of the [Cancer Biology Commons](#), [Computational Biology Commons](#), [Discrete Mathematics and Combinatorics Commons](#), [Dynamic Systems Commons](#), [Numerical Analysis and Computation Commons](#), and the [Ordinary Differential Equations and Applied Dynamics Commons](#)

Recommended Citation

Scott, Paul, "Dynamics of Gene Networks in Cancer Research" (2017). *Electronic Theses and Dissertations*. 1538.
<https://digitalcommons.georgiasouthern.edu/etd/1538>

This thesis (open access) is brought to you for free and open access by the Graduate Studies, Jack N. Averitt College of at Digital Commons@Georgia Southern. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.

DYNAMICS OF GENE NETWORKS IN CANCER RESEARCH

by

PAUL AARON SCOTT

(Under the Direction of Hua Wang)

ABSTRACT

Cancer prevention treatments are being researched to see if an optimized treatment schedule would decrease the likelihood of a person being diagnosed with cancer. To do this we are looking at genes involved in the cell cycle and how they interact with one another. Through each gene expression during the life of a normal cell we get an understanding of the gene interactions and test these against those of a cancerous cell. First we construct a simplified network model of the normal gene network. Once we have this model we translate it into a transition matrix and force changes on it. Observing the effects of the changes we see the interactions each gene has with other genes within the network. Using the observed interactions we construct a set of differential equations that represent the network dynamics. Using numerical methods and the rough system of equations, we find an approximated system of equations that accurately predicts the dynamics of the normal gene network.

Key Words: Dynamics, Biology, Gene networks, Differential equations, Dynamical systems

2009 Mathematics Subject Classification: 92B05, 90C90

DYNAMICS OF GENE NETWORKS IN CANCER RESEARCH

by

PAUL AARON SCOTT

B.S., Gordon State College, 2014

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial
Fulfillment
of the Requirement for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA

©2017

PAUL AARON SCOTT

All Rights Reserved

DYNAMICS OF GENE NETWORKS IN CANCER RESEARCH

by

PAUL AARON SCOTT

Major Professor: Hua Wang

Committee: Yan Wu
Guang Peng

Electronic Version Approved:

May 2017

DEDICATION

To my dad, who always pushed me to do better.

ACKNOWLEDGMENTS

I wish to acknowledge Dr. Wang and Dr. Wu for their patience with me while writing this thesis. My family for their support throughout the entire learning process and their ability to listen. I would also like to thank my best friend, James, for always knowing when to lure me away from my work to take a break.

TABLE OF CONTENTS

	Page
DEDICATION	2
ACKNOWLEDGMENTS	3
LIST OF FIGURES	7
CHAPTER	
1 Introduction	9
1.1 Biological Background	9
1.2 Graphs, Random Walk, and Stationary Distribution	11
1.2.1 Basic Graph Theory	11
1.2.2 Random Walks	13
1.2.3 The Transition Matrix and Stationary Distribution	14
1.3 Systems of Differential Equations	14
1.4 Our Approach	16
2 Forced “Evolutions” on Genes	18
2.1 Initial Set Up	18
2.2 Modeling the Changes	20
2.3 Results	23
3 The Network’s Self-Evolution	29
3.1 Data gathering	29
3.1.1 Natural Evolution of the Network	29

	5
3.1.2 Initial setup of system of differential equations	30
3.1.3 Generating the data for use in finding the parameters	30
3.2 Finding parameters of the equations	31
3.2.1 First set, an example	31
3.2.2 First matrices set up	32
3.3 Second set	33
3.3.1 Second Least-Squares Equation Work	33
3.3.2 Second matrices set up	34
3.4 Third set	34
3.4.1 Third Least-Squares Equation Work	34
3.4.2 Third matrices set up	35
3.5 Fourth set	36
3.5.1 Fourth Least-Squares Equation Work	36
3.5.2 Fourth matrices set up	36
3.6 Fifth set	37
3.6.1 Fifth Least-Squares Equation Work	37
3.6.2 Fifth matrices set up	38
3.7 Gauss-Seidel Method	38
3.8 Results	40
3.8.1 Plots of the evolution of the network	40
3.8.2 The comparison plots that show the observed dynamics with the calculated dynamics of the system	40

	6
3.8.3 The System of Differential Equations Found	41
3.8.4 Stability Analysis	42
4 In Closing	49
REFERENCES	50

LIST OF FIGURES

Figure		Page
1.1	Cell Cycle (https://oncogenesandcancer.wordpress.com/)	10
1.2	Vertex and edges of a graph (http://world.mathigon.org/)	11
1.3	An undirected graph (http://i.stack.imgur.com/uNgXD.png)	12
1.4	An electrical shunt circuit showing the pathways available for an electron to go down and the corresponding graph (http://www.scielo.br/) . . .	12
1.5	The adjacency matrix and the transition matrix built from the graph (http://image.slidesharecdn.com/)	15
2.1	A gene network.	18
2.2	The network with artificial nodes.	19
2.3	The time response for x_1 when forcing a change on its corresponding dependent variables.	24
2.4	The time response for the x_2 variable when forcing a change on its corresponding dependent variables.	25
2.5	The time response for the x_3 variable when forcing a change on its corresponding dependent variables.	26
2.6	The time response for the x_4 variable when forcing a change on its corresponding dependent variables.	27
2.7	The time response for the x_5 variable when forcing a change on its corresponding dependent variables.	28
3.1	State of X_1 and its finite difference.	40
3.2	State of X_2 and its finite difference.	41

3.3	State of X_3 and its finite difference.	41
3.4	State of X_4 and its finite difference.	42
3.5	State of X_5 and its finite difference.	42
3.6	Actual finite differences for X_1 and the finite differences estimated from the parameters found using least squares methods.	43
3.7	Actual finite differences for X_2 and the finite differences estimated from the parameters found using least squares methods.	44
3.8	Actual finite differences for X_3 and the finite differences estimated from the parameters found using least squares methods	45
3.9	Actual finite differences for X_4 and the finite differences estimated from the parameters found using least squares methods	46
3.10	Actual finite differences for X_5 and the finite differences estimated from the parameters found using least squares methods	46
3.11	Actual state of X_1 and the predicted state found using the RK-4 method.	47
3.12	Actual state of X_2 and the predicted state found using the RK-4 method.	47
3.13	Actual state of X_3 and the predicted state found using the RK-4 method.	47
3.14	Actual state of X_4 and the predicted state found using the RK-4 method.	48
3.15	Actual state of X_5 and the predicted state found using the RK-4 method.	48

CHAPTER 1

INTRODUCTION

1.1 Biological Background

The cell cycle is what biological cells go through from their point of creation to either cell death, a stable existence, or cell division. This cycle can be broken down into two large phases: Interphase and Mitotic (M) phase. Interphase is then broken down into several other phases with the first being the G1 phase. The G1 phase of interphase is characterized by cell growth. From G1 phase the cell enters S phase. During this phase the cell has reached an optimal size and is now synthesizing a copy of its DNA within its nucleus as well as setting the stage for G2 phase. Once it enters G2 phase from S phase the cell begins to grow more as well as make more organelles and proteins while splitting its contents between two sides of the cell. When G2 phase ends, interphase is now over and the cell moves into M phase, which is where the cell begins to divide into two new cells. From this point the cells that are made will either enter interphase or enter an alternative phase known as G0 phase. During G0 phase the cell is in a steady state and is just performing its duties. It may or may not enter G1 phase at a later time. Depending on what type of cell it is the cell may or may not go through every phase of the cell cycle. Figure 1.1 shows the diagram of a typical cell cycle and where cell cycle checkpoints are located within that cycle.

Roughly speaking, cancer occurs when cells do not behave as normal cells do. This is usually due to the fact that there has been some mistake that affects this cell cycle. There are three defects within the cell cycle that can cause cancer. Of these there are unscheduled proliferation, genomic instability, and chromosomal instability. All of these defects can be contributed to cyclin-dependent kinases (CDKs) which are used to control the cell cycle through individual check points between cell phases. These CDKs, as well as several other types of proteins, cause unchecked progression through the cell cycle when unregulated

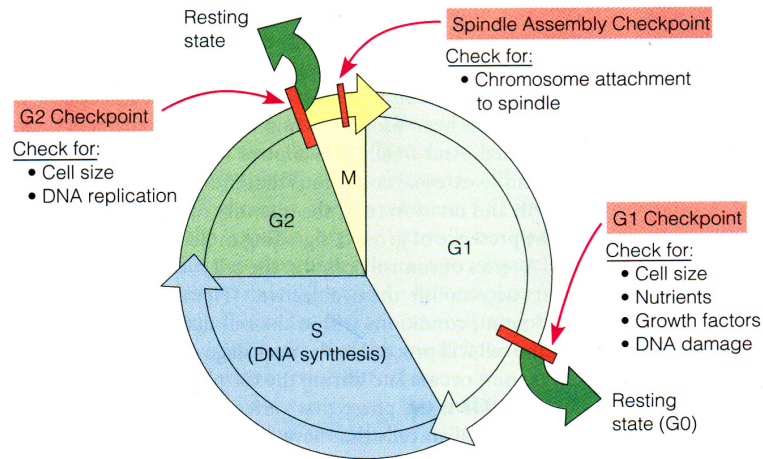


Figure 1.1: Cell Cycle (<https://oncogenesandcancer.wordpress.com/>)

[1].

As mentioned above CDKs can wreak havoc on a cell if their activity is not controlled. Unscheduled cell proliferation can happen because of misregulated CDKs as well as chromosomal instabilities [1]. Unscheduled proliferation happens when a cell begins to go through the cell division process without any of the cell cycle check points preventing it from going through with cell division. This results in a runaway process in which one could get a multitude of cells, densely packed in a small area.

Genomic instability occurs when there is a high frequency of mutations within a cell's genome which can in turn mean changes in gene sequences, chromosomal rearrangements and sometimes aneuploidy. Changes in gene sequences could cause the proteins that are to be made by certain genes to be inert and not work, thus causing a lack of what could potentially be a protein that is essential to the cell cycle. Chromosomal rearrangements could mean that some genes do not get to be produced because of their physical location on a chromosome prevents their replication from happening.

Chromosomal instability occurs when a chromosome or multiple chromosomes are reproduced too many times so that a cell that has just completed cell division now had 3 or

more copies of a chromosome or a chromosome or parts of a chromosome get deleted all together.

The genes that are associated with this cell cycle check point system can be represented as a network by using a graph. The graphs that can represent these gene networks are constructed with vertices representing the genes and directed edges going to or from each vertex representing the interaction paths between each gene. We shall talk about this next.

1.2 Graphs, Random Walk, and Stationary Distribution

1.2.1 Basic Graph Theory

Mathematically speaking a graph is simply a diagram representing a system of connections/relations between two or more objects which are represented as dots. Each dot within a system is a vertex and the connections between pairs of vertices are called edges. An example is shown in Figure 1.2 with a vertex and an edge labeled.

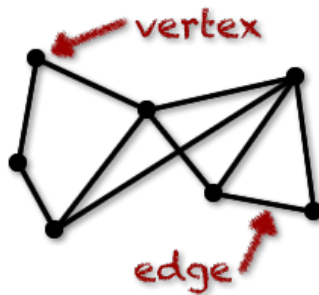


Figure 1.2: Vertex and edges of a graph (<http://world.mathigon.org/>)

Graphs can be undirected, which means that the giving/receiving ends of a relationship between two vertices are not specified. For instance, if each vertex represented a different place around a town and each edge represented a road to one place from another then a person who is at place 1 could go to place 5, and using the same path, go back to place 1.

This would be an undirected edge in the graph in Figure 1.3.

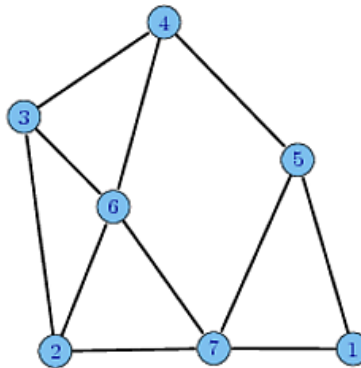


Figure 1.3: An undirected graph (<http://i.stack.imgur.com/uNgXD.png>)

Graphs can be directed as well. This means that each vertex could have edges going to or leaving it. With these types of graphs movement around a graph is restricted in a sense. For example, we could model an electrical circuit as a directed graph, with a power source and resistors in parallel to one another. Within this circuit an electron can only go from the negative pole of the power source to the positive pole of the power source by moving further away from the negative pole without back-tracking towards it, see Figure 1.4.

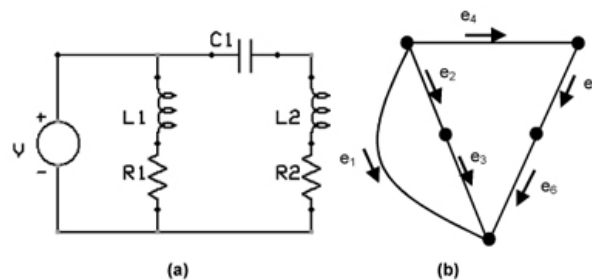


Figure 5. (a) Shunt network; (b) associated directed graph.

Figure 1.4: An electrical shunt circuit showing the pathways available for an electron to go down and the corresponding graph (<http://www.scielo.br/>)

In both directed and undirected graphs, movement around the graph can be modeled through mathematical tools which will be discussed in the next few sections.

1.2.2 Random Walks

A specific type of movement around a graph can be classified as a random walk. The random walk is a process in which, while going around the graph, the edge of which one goes from one particular vertex to another is chosen at random. To exemplify this is a version of the famous “Drunkards Walk.” Using Figure 1.3 above as an example the “Drunkards Walk” uses a scenario similar to the above mentioned town with numerous places and a man, who would not be driving but walking for the safety of him and the public, who drunkenly goes from one place to another with no tact or reason behind going other than he has forgotten where he was going and does not remember where he has been. Keeping in mind he has no memory of where he has been and no idea where he is going he wanders from one place to another by choosing to go to a new place at a random. This wondering around aimlessly shows the random walk where the places are vertices and the routes he takes to get to each one are the edges of the graph he is traversing.

You can have random walks on both undirected and directed graphs. The above example of a random walk was using an undirected graph. To imagine a random walk on a directed graph think about the electrical circuit modeled as a graph in Figure 1.4. Each time an electron goes through the circuit it has several different ways it can go depending on the resistance levels of each path it can take. For now we shall assume the resistance of each pathway is the same so that there is an equal probability of an electron going through a particular resistor each time. With all that said if an electron leaves the negative pole of the power source then from the power source to the junction before the parallel resistors the electron is on one path. Once it gets to the junction it will take one path in particular through a resistor to the junction after the parallel resistors and then to the positive pole of the power source. Then the next electron that goes through will follow a similar path although it could possibly go through a different resistor than the previous electron. This shows how, in a directed graph, things can move at random and create the random walk in

a directed graph.

1.2.3 The Transition Matrix and Stationary Distribution

The directed graphs that model certain systems can also be represented using matrices. A matrix that represents a graph can be built by making a square matrix, where each column represents incoming interactions (edges) between a particular vertex and the other vertices of that graph and each row represents the outgoing interactions between a particular vertex and the other vertices of that graph. The matrix will be square because the number of columns must match the number of rows, which is the number of vertices in your graph. To start, where there is an interaction between two vertices there will be an entry of 1. Where there is no interaction between two vertices a zero is normally placed in the matrix representation. This matrix which just indicates which vertices a vertex is connected to is called an adjacency matrix.

We can construct from the adjacency matrix the transition matrix. To do this each nonzero entry within a row is divided by the number of nonzero entries from that particular row. When this is done, the new entries in each row sum up to one. The matrix described is called a transition matrix and it helps determine the weights of each edge of the graph based on a set of initial conditions and how many times the graph is introduced to new conditions. An example of this representation and the graph the matrix represents is shown in Figure 1.5.

The stationary distribution is basically the final distribution over the vertices of a graph. It can be found using the transition matrix of a graph by finding the eigenvector corresponding to the eigenvalue of one. We shall see the use of such stationary distributions throughout this thesis.

1.3 Systems of Differential Equations

A differential equation describes the relation of a given variable with derivatives of

Adjacency and Transition Matrix

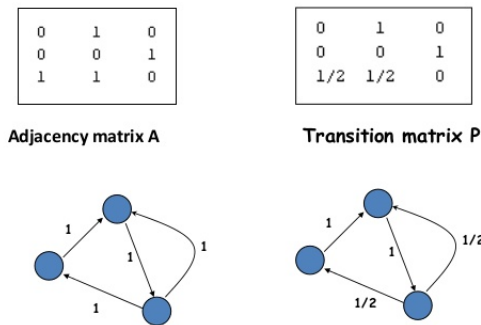


Figure 1.5: The adjacency matrix and the transition matrix built from the graph (<http://image.slidesharecdn.com/>)

several orders of itself and possibly other variables. Given the nature of a differential equation it can be used to model how a certain function or physical object behaves. Differential equations have been used in many fields including physics, engineering, biology and economics.

Sometimes it is necessary to use systems of differential equations to model and describe the behaviors of a real object. For example a first order system of equations usually involve first order differential equations of several variables where each differential equation represents the interaction between the variable the equation is modeling and those corresponding variables. The first order system of differential equations describes the rates of changes of the specified variable by using its corresponding interactions with other variables.

Suppose that you have a physical system where each component of that system had a reaction based on input from other components within that same system and you needed to model the reaction of each component. How could you do that? If the system only requires that you show the rate of change of reaction of each component then you could

create a system of differential equations where each equation was a first order differential with respect to time that was equal to the sum of reactions it had to each of the components of that system which interacted with it first-hand.

How would you know how each component reacted with the others within that system to create the differential equations? If you have a diagram of that particular system which showed the components, inputs, and outputs then you could have a way to find the differential equations involved in that system. You can take the diagram and translate it into a directed graph and from there to a transition matrix. Using the transition matrix you could use the columns that represent the incoming edges of the graph to a particular vertex to find which vertex (component) interacted with it and create a rough differential equation for that vertex and find the other rough equations similarly. Then once one has these they can then use computational tools to find the exact ways each component reacts to each other component and refine the differential equation for each component to where they have an accurate representation for the change in reaction for each component. There will be a more in depth look at how this process is done and the numerical methods involved further in the discussion of the research at hand.

1.4 Our Approach

This research will focus on modeling the gene expressions and interactions of a selected few CDKs and proteins with each other through the cell cycle in the form of a network. Our motive for doing so is to see if preventative treatments of Tamoxifen or other drugs could be administered on an optimized schedule to minimize the chance of a patient from getting breast cancer [3, 4, 5].

In order to create models for the dynamics of a gene network we will first use a directed graph of the interactions of CDKs and proteins relevant to specific check points within the cell cycle and remove all but the most necessary CDKs and proteins from the graph. Each vertex of the graph shall represent a gene expression of a CDK or other protein

and each edge of this graph shall show interaction between two associated vertices.

Once we have a gene network that we feel most appropriately fits the scope of our research, we will create an adjacency matrix that represents the interactions (edges) between each vertex with ones and zeros where there is no interaction between vertices present.

We will then use this matrix to create a stochastic (probability or transition) matrix where all values are between zero and one and each row adds up to one, using a Matlab program. Each column within this matrix shall represent output edges from a single vertex to other vertices and each row shall represent input edges from other vertices to a single vertex.

We will use this transition matrix to determine the effects of the change in one vertex output on other vertices for each vertex. Once we have modeled the effect each vertex has on the others we shall use the data to create a set of general differential equations to model the dynamics of each vertex within this system. We will refine the equations by performing a parameter search by both least squares method and Gauss-Seidel method.

CHAPTER 2

FORCED “EVOLUTIONS” ON GENES

2.1 Initial Set Up

To begin this project we obtained a list of genes, represented by the proteins they code for, which play a vital part of the cell cycle check point system. Of these genes we came up with 5 key genes that are given and shown in Figure 2.1. Once we had this graph we then constructed a new graph with two artificial nodes for modeling purposes. The artificial nodes will help create useful random walks as one node acts as a transition node for which all the real, gene specific nodes, flow to and the other that acts as a restarting node to return the flow back to the nodes to reset the cycle as shown in Figure 2.2.

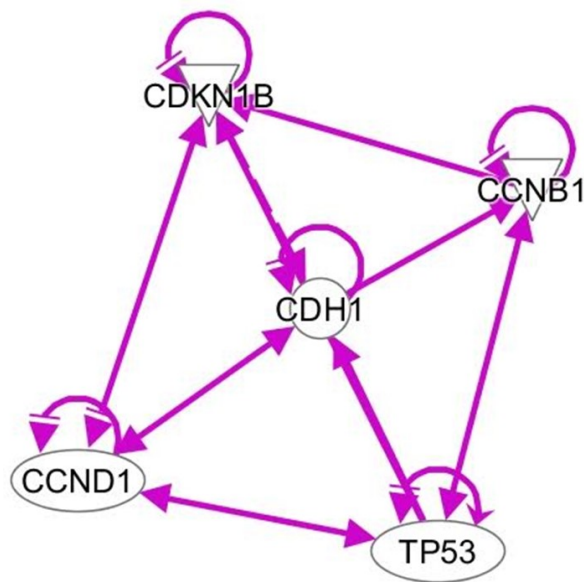


Figure 2.1: A gene network.

Utilizing the relationships indicated by the graph shown we made the following 7×7 adjacency matrix:

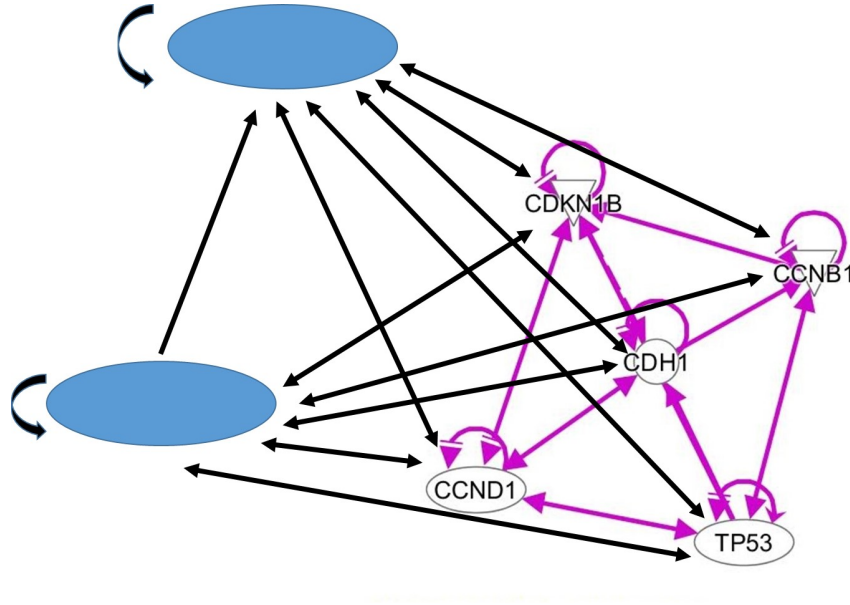


Figure 2.2: The network with artificial nodes.

$$\begin{bmatrix}
 & R & X_1 & X_2 & X_3 & X_4 & X_5 & T \\
 R & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\
 X_1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
 X_2 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\
 X_3 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
 X_4 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\
 X_5 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\
 T & 1 & 0 & 0 & 0 & 0 & 0 & 1
 \end{bmatrix}.$$

Then using a script written in MATLAB, we called this adjacency matrix from the excel file and constructed a transition matrix. In order to construct the said transition matrix we divided each entry by the number of nonzero elements contained within the corresponding row as we did for the first row here:

$$\frac{1}{\text{Number of Non Zero Terms}} = \frac{1}{6}.$$

Once we obtained this number for each row, each between zero and one, we replace each nonzero element with this new number so that each row added up to one, as shown below. We then added this transition matrix to the excel file we used to store the adjacency matrix in a different spreadsheet. Once we have this transition matrix:

$$\begin{bmatrix} & R & X_1 & X_2 & X_3 & X_4 & X_5 & T \\ R & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\ X_1 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & 0 & \frac{1}{5} & 0 & \frac{1}{5} \\ X_2 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} \\ X_3 & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ X_4 & \frac{1}{5} & 0 & 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ X_5 & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ T & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix},$$

we can then use this matrix to find the behavior of the network.

2.2 Modeling the Changes

In order to get an idea of the behavior of the network dynamics we had to decide how the transition matrix would change with a given set of conditions. To do this, we created another MATLAB script which would model the change of variable to variable interactions.

In order to model these interactions the script used had to perform the following.

- The column representing the independent variable would be multiplied by a range of scalar numbers in a while-loop.
- The nonzero numbers that were not in that specified column would be altered so that those nonzero numbers plus the elements in the specified column would sum up one.
- Take the stationary distribution of the specified variable that we were looking to model changes after and place it in a vector that would record all of the eigenvector

values for the specified variable to variable interaction.

For example if you take the second column to be the independent variable then you will multiply that column by a scalar, in our case 0.2:

$$x_1 = \begin{pmatrix} 0.2 \end{pmatrix} \begin{pmatrix} \frac{1}{6} \\ \frac{1}{5} \\ \frac{1}{6} \\ 0 \\ 0 \\ \frac{1}{6} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{30} \\ \frac{1}{25} \\ \frac{1}{30} \\ 0 \\ 0 \\ \frac{1}{30} \\ 0 \end{pmatrix}$$

Once we have obtained the new values for this column of the transition matrix we replace the old column within the matrix with this one:

$$\begin{bmatrix} & R & X_1 & X_2 & X_3 & X_4 & X_5 & T \\ R & \frac{1}{6} & \frac{1}{30} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\ X_1 & \frac{1}{5} & \frac{1}{25} & \frac{1}{5} & 0 & \frac{1}{5} & 0 & \frac{1}{5} \\ X_2 & \frac{1}{6} & \frac{1}{30} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} \\ X_3 & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ X_4 & \frac{1}{5} & 0 & 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ X_5 & \frac{1}{6} & \frac{1}{30} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ T & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

If you will notice, however, now the rows do not individually sum up to one. We now need to adjust all other non-zero entries within each row to compensate for the new numbers in the second column so that the rows sum up to one again. In order to do that we look at the second column as see which row entries within that column have non-zero entries. These will be the rows that need adjusting. Once we have made a note of these

rows we begin to subtract each non-zero new number from the second column from one and divide each of these rows by the number of non-zero entries in each row minus one to account for the new second column entries which will stay the same. We do this as shown in the next example with the first row:

$$\frac{1 - \text{Multiple of Specified Vector Element}}{\text{Number of Non Zero Terms} - 1} = \frac{1 - \frac{1}{30}}{5} = \frac{29}{150}.$$

Once we have this we replace all non-zero terms within the row, with the exception of the second column value, with this new number so:

$$[R]_1 = \left(\frac{29}{150} \quad \frac{1}{30} \quad \frac{29}{150} \quad \frac{29}{150} \quad \frac{29}{150} \quad \frac{29}{150} \quad 0 \right).$$

Likewise, you do this for the rest of the rows until you get the completed matrix:

$$\begin{bmatrix} R & X_1 & X_2 & X_3 & X_4 & X_5 & T \\ R & \frac{29}{150} & \frac{1}{30} & \frac{29}{150} & \frac{29}{150} & \frac{29}{150} & \frac{29}{150} & 0 \\ X_1 & \frac{6}{25} & \frac{1}{25} & \frac{6}{25} & 0 & \frac{6}{25} & 0 & \frac{6}{25} \\ X_2 & \frac{29}{150} & \frac{1}{30} & \frac{29}{150} & \frac{29}{150} & 0 & \frac{29}{150} & \frac{29}{150} \\ X_3 & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ X_4 & \frac{1}{5} & 0 & 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ X_5 & \frac{29}{150} & \frac{1}{30} & 0 & \frac{29}{150} & \frac{29}{150} & \frac{29}{150} & \frac{29}{150} \\ T & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

This is your new transition matrix. Once you have it, you take the stationary distribution corresponding to the new transition matrix and repeat the procedure on the new transition matrix with a new scalar until one of the following conditions are reached:

- An entry within the transition matrix is less than zero.
- An entry within the transition matrix is greater than one.

- The maximum number of scalars from which to multiply the transition matrix with has been reached.

2.3 Results

The following plots represent how each variable reacted to the other relevant variables. The x -axis of each plot represents a step in time while the y -axis represents the change in the variable. The change in the variable, which can be defined as an approximation of the derivative, was found using the forward, central, and backward difference methods. These differences can be used to approximate a differential equation when the differences are plotted against time.

We first show an example of using the central difference method by taking a set of data recorded in the form of a matrix:

$$\begin{bmatrix} Time & Value \\ 1 & 8 \\ 2 & 9 \\ 3 & 4 \\ 4 & 3 \end{bmatrix}$$

Once you have data then you can find the forward difference, which is the difference between the first and second data value divided by the time step between the two:

$$\text{Forward Difference} = \frac{9 - 8}{1}$$

Then to find the central difference we find the difference between the first and third data value and divide that by the time step between the two numbers:

$$\text{First Central Difference} = \frac{4 - 8}{2}$$

$$\text{Second Central Difference} = \frac{3 - 9}{2}$$

And to find the backward difference we subtract the second to last data value from the last data value and divide that by the time step between the two values.

$$\text{Backward Difference} = \frac{3 - 4}{1}$$

After doing this we have a central difference vector which we can use as an approximation for the change of the data, or the derivative, from one point of data to the next:

Time	ΔValue
1	1
2	-2
3	-3
4	-1

Using data gathered in a manner similar to the above example we then plotted the time response of each of the variables' change in flow for each of these variables as shown in the resulting figures below:

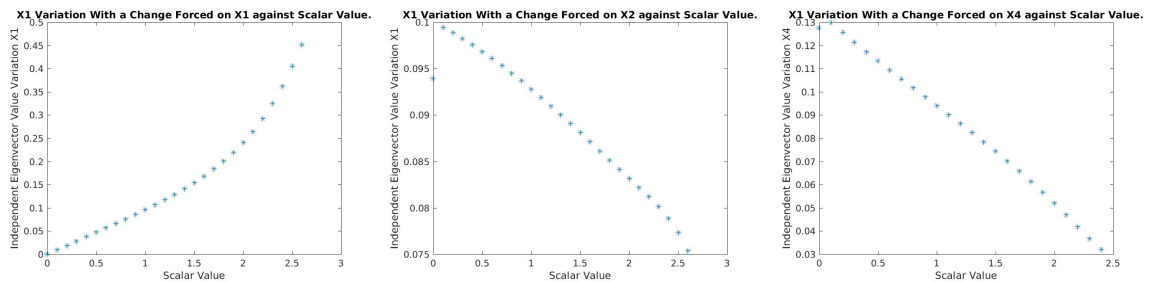


Figure 2.3: The time response for x_1 when forcing a change on its corresponding dependent variables.

As you can see in Figure 2.3, the first plot shows the rate of change of the eigenvector value corresponding to x_1 when a change is forced onto itself. This shows that the rate of change corresponding to x_1 to a change forced on itself is concave up and increasing. With the second plot in the figure, which shows the rate of change of x_1 when a change is forced on x_2 , is concave down and decreasing. Likewise in the third plot in the figure, the rate of change that occurs when x_4 has a changed forced on it is linear and decreasing.

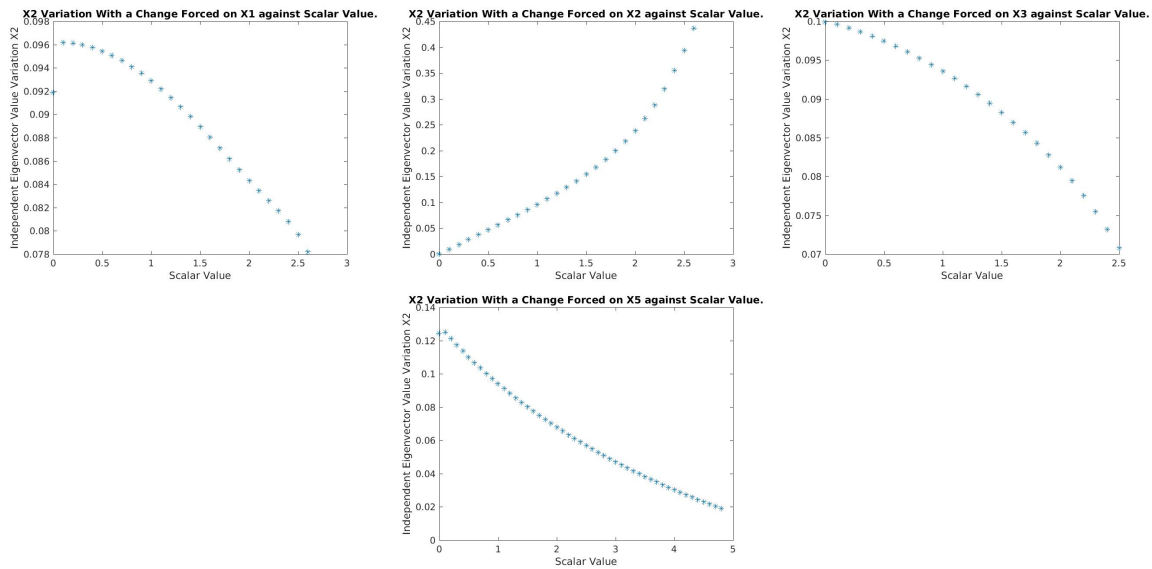


Figure 2.4: The time response for the x_2 variable when forcing a change on its corresponding dependent variables.

In Figure 2.4, the first plot shows the rate of change of the eigenvector value corresponding to x_2 when a change is forced onto x_1 . This shows that the rate of change corresponding to x_2 when a change is forced on x_1 is concave down and decreasing. With the second plot in the figure, which shows the rate of change of x_2 when a change is forced on itself, is concave up and increasing. Likewise in the third plot in the figure, the rate of change that occurs when x_3 has a changed forced on it is concave down and decreasing. The last plot shows the change when a change is forced on x_5 . That plot is concave up and decreasing.

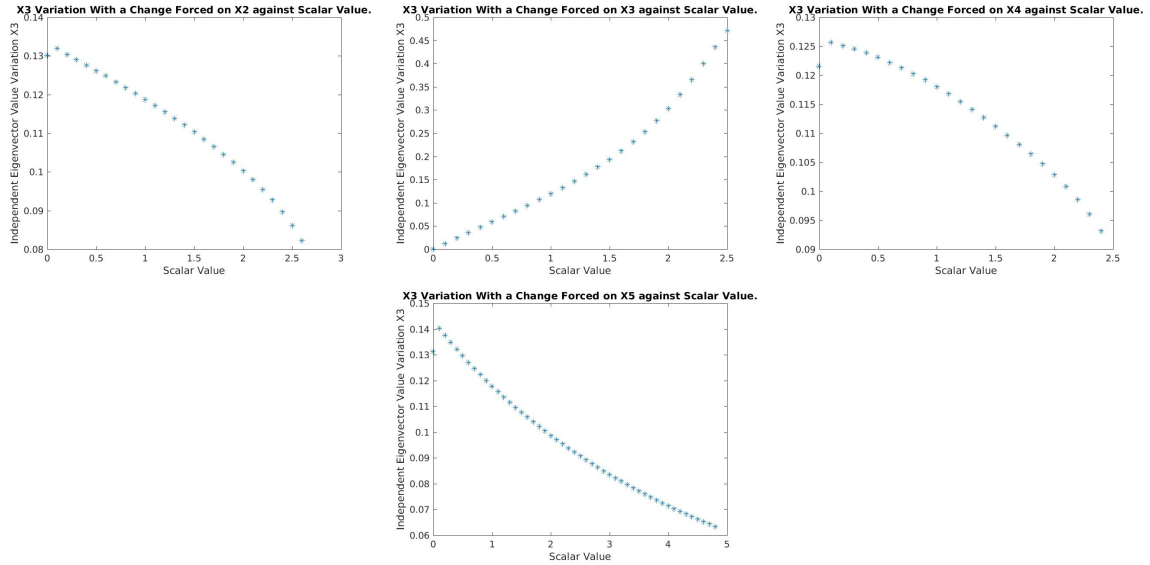


Figure 2.5: The time response for the x_3 variable when forcing a change on its corresponding dependent variables.

The first of these plots, in Figure 2.5, shows the change in x_3 when a change is forced upon x_2 . The plot shows that when this variable is changed it reduces the eigenvector value corresponding to x_3 resulting in a concave down and decreasing plot. When x_3 is plotted against changes of itself it results in a concave up and positive plot. The third plot shows the rate of change of x_4 when x_4 has a change forced on it. This plot shows a concave down and decreasing trend. When finding how x_5 effects x_3 , we found this concave up and decreasing.

Almost all of the plots, in Figure 2.6, that show change in x_4 all show some linearity. The first of these plots shows the change in x_3 when a change is forced upon x_4 . The plot shows that when this variable is changed it reduces the eigenvector value corresponding to x_3 resulting in a linear, decreasing trending plot. The second plot shows the rate of change of x_3 when x_4 has a change forced upon it. This plot shows a concave up, increasing trend. The third plot shows the decreasing and linear trend associated with the changes forced upon (x_5).

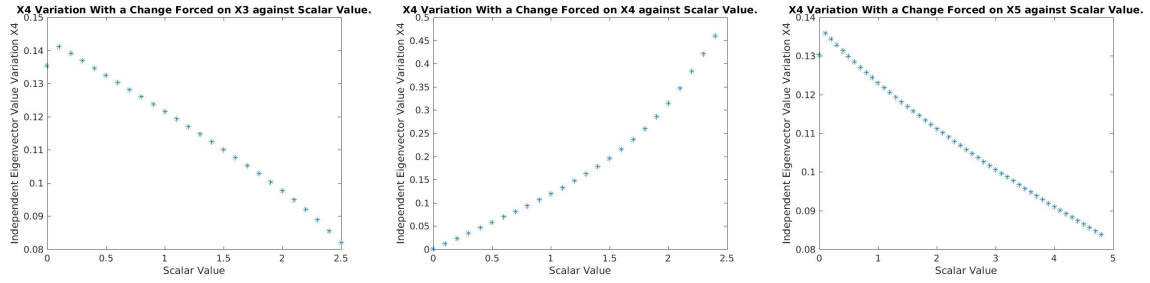


Figure 2.6: The time response for the x_4 variable when forcing a change on its corresponding dependent variables.

The first of the plots, in Figure 2.7, shows the change in x_5 when a change is forced upon x_1 which is concave down with a decreasing trend. The second plot shows the rate of change of x_4 when x_3 has a change forced upon it. This plot shows a linear decreasing trend. The third plot shows the concave down, decreasing trend associated with the changes forced upon (x_4). When finding how x_5 effects x_5 , we found this concave down, positively sloped plot.

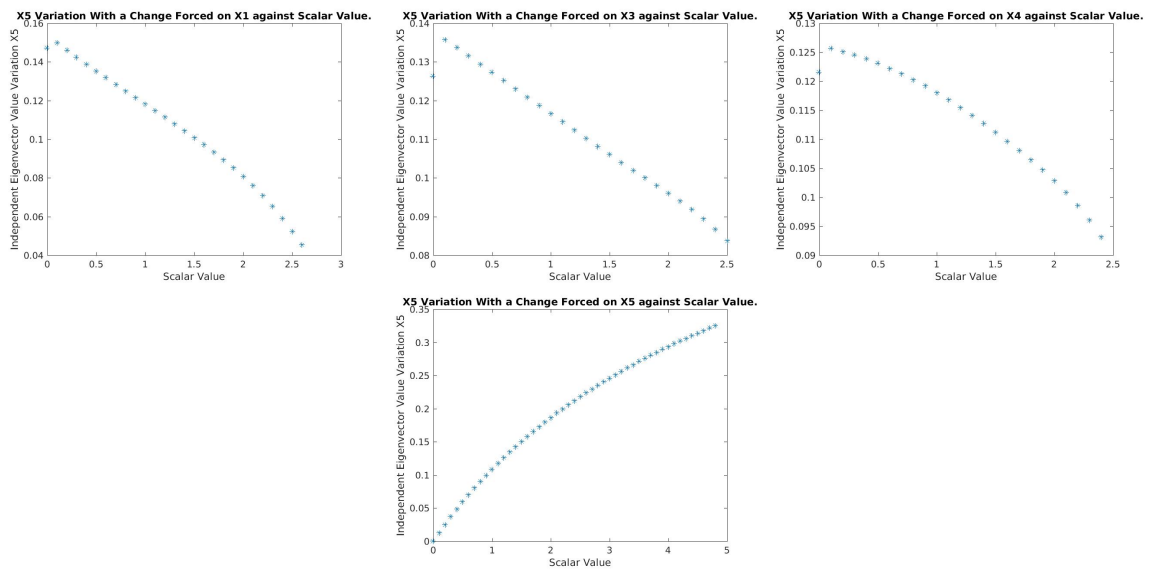


Figure 2.7: The time response for the x_5 variable when forcing a change on its corresponding dependent variables.

CHAPTER 3

THE NETWORK'S SELF-EVOLUTION

3.1 Data gathering

After getting a general idea of how the network changed with a change in one variable we then advanced to modeling the entire gene network.

3.1.1 Natural Evolution of the Network

To start doing this we initially need data from the network as it evolved naturally through time, similarly to how we did for the first part of this research but with a few changes. In order to obtain this data from the network we would first pull the original transition matrix from the excel file, the same one that is used in the first part of our research. Once the Excel file is read into the MATLAB program we then multiply a initial and arbitrary probability distribution, in our case:

$$X^0 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

by the transition matrix:

$$\begin{bmatrix} & R & X_1 & X_2 & X_3 & X_4 & X_5 & T \\ R & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\ X_1 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & 0 & \frac{1}{5} & 0 & \frac{1}{5} \\ X_2 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} \\ X_3 & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ X_4 & \frac{1}{5} & 0 & 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ X_5 & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ T & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

Once this is done we now have a new probability distribution:

$$X^1 = \begin{bmatrix} \frac{1}{5} & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{5} & 0 & \frac{1}{5} \end{bmatrix},$$

which we can multiply to the transition matrix and come up with a new probability distribution. We repeat this loop until the values in each element in the are the same each loop. Once we have these values we can then find the finite differences, as discussed in Chapter 2, of the data and store that in an excel file.

3.1.2 Initial setup of system of differential equations

After creating these new plots we began the process of making generic forms of the differential equations that would represent the dynamics of each variable within the system. Based on the plots we got the following five differential equations:

$$\dot{x}_1 = r_{11}x_1 + r_{12}x_1x_2 + r_{13}x_5$$

$$\dot{x}_2 = r_{21}x_1 + r_{22}x_2 + r_{23}x_3 + r_{24}x_1x_2$$

$$\dot{x}_3 = r_{31}x_2 + r_{32}x_3 + r_{33}x_4 + r_{34}x_5 + r_{35}x_3^2$$

$$\dot{x}_4 = r_{41}x_1 + r_{42}x_1x_3 + r_{43}x_4 + r_{44}x_5$$

$$\dot{x}_5 = r_{51}x_2x_3 + r_{52}x_4 + r_{53}x_5$$

Once these differential equations were formed we then moved on to finding the right values to use for the parameters r_{ij} in the equations.

3.1.3 Generating the data for use in finding the parameters

In order to find the parameters of the equations we came up with we needed to first have some numerical data so that we can find precisely what each parameter is. To do this we created another MATLAB script that reads the transition matrix from the excel file then,

using a for-loop, we multiplied the transition matrix by an initial probability distribution $[0 \ 1 \ 0 \ 0 \ 0 \ 0]$ to get a new vector and repeated this for a few hundred times to collect the samples. In doing this we were able to find what the transition matrix's probability distribution converged to, which was then considered the stationary distribution. Using the data provided from doing this we could now proceed to estimate the parameters of the system of differential equations.

3.2 Finding parameters of the equations

3.2.1 First set, an example

In order to estimate the parameters of the system of equations we first had to work out a few things. First we decided that based on the size of the system using a method such as Monte Carlo simulation was not needed and that we could use a Least-Squares method to estimate the likelihood of the parameters. To do this we first needed to derive the system of equations for the parameters based on the method of least-squares, as shown with the work done below for the first differential equation out of the system:

First we needed to set up the performance index.

$$J = \frac{1}{2} \sum_{i=1}^N [r_{11}x_{1i} + r_{12}x_{2i} + r_{13}x_{5i} - \dot{x}_{1i}]^2$$

,

where \dot{x}_{1i} are the derivative samples.

After finding the general least-squares equation we find the derivative of each, with respect to a particular parameter in the following fashion. Each of these derivatives with respect to a parameter are set to zero.

$$\begin{aligned}
\frac{\partial J}{\partial r_{11}} &= r_{11} \sum_{i=1}^N x_{1i}^2 + r_{12} \sum_{i=1}^N x_{1i}x_{2i} + r_{13} \sum_{i=1}^N x_{1i}x_{5i} - \sum_{i=1}^N x_{1i}\dot{x}_{1i} = 0 \\
\frac{\partial J}{\partial r_{12}} &= r_{11} \sum_{i=1}^N x_{1i}x_{2i} + r_{12} \sum_{i=1}^N x_{2i}^2 + r_{13} \sum_{i=1}^N x_{2i}x_{5i} - \sum_{i=1}^N x_{2i}\dot{x}_{1i} = 0 \\
\frac{\partial J}{\partial r_{13}} &= r_{11} \sum_{i=1}^N x_{1i}x_{5i} + r_{12} \sum_{i=1}^N x_{2i}x_{5i} + r_{13} \sum_{i=1}^N x_{5i}^2 - \sum_{i=1}^N x_{5i}\dot{x}_{1i} = 0
\end{aligned}$$

Therefore, the parameters satisfy the following system of equations:

$$\begin{aligned}
\sum_{i=1}^N x_{1i}\dot{x}_{1i} &= r_{11} \sum_{i=1}^N x_{1i}^2 + r_{12} \sum_{i=1}^N x_{1i}x_{2i} + r_{13} \sum_{i=1}^N x_{1i}x_{5i} \\
\sum_{i=1}^N x_{2i}\dot{x}_{1i} &= r_{11} \sum_{i=1}^N x_{1i}x_{2i} + r_{12} \sum_{i=1}^N x_{2i}^2 + r_{13} \sum_{i=1}^N x_{2i}x_{5i} \\
\sum_{i=1}^N x_{5i}\dot{x}_{1i} &= r_{11} \sum_{i=1}^N x_{1i}x_{5i} + r_{12} \sum_{i=1}^N x_{2i}x_{5i} + r_{13} \sum_{i=1}^N x_{5i}^2
\end{aligned}$$

3.2.2 First matrices set up

If you examine the system of equations given above you will notice that it can also be written in the matrix form as shown below, however, instead of the x -values being in a column and being multiplied by parameters it is reversed.

$$\begin{pmatrix} \sum_{i=1}^N x_{1i}^2 & \sum_{i=1}^N x_{2i} & \sum_{i=1}^N x_{1i}x_{5i} \\ \sum_{i=1}^N x_{1i}x_{2i} & \sum_{i=1}^N x_{2i}^2 & \sum_{i=1}^N x_{2i}x_{5i} \\ \sum_{i=1}^N x_{1i}x_{5i} & \sum_{i=1}^N x_{2i}x_{5i} & \sum_{i=1}^N x_{5i}^2 \end{pmatrix} \begin{pmatrix} r_{11} \\ r_{12} \\ r_{13} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N x_{1i}\dot{x}_{1i} \\ \sum_{i=1}^N x_{2i}\dot{x}_{1i} \\ \sum_{i=1}^N x_{5i}\dot{x}_{1i} \end{pmatrix}.$$

Once we have the least-squares equation in matrix form we can now use a MATLAB script to compute the correct parameters. Once these parameters are found they are stored in a “.dat” file to be used in the next stage of this research.

3.3 Second set

This set is done similarly to the above example.

3.3.1 Second Least-Squares Equation Work

The performance index:

$$J = \frac{1}{2} \sum_{i=1}^N [r_{21}x_{1i} + r_{22}x_{2i} + r_{23}x_{3i} + r_{24}x_{1i}x_2 - \dot{x}_{2i}]^2$$

Similar to the previous case, by setting the derivatives of this parameter index to zero, we find the following system of equations:

$$\begin{aligned} \frac{\partial J}{\partial r_{21}} &= r_{21} \sum_{i=1}^N x_{1i}^2 + r_{22} \sum_{i=1}^N x_{1i}x_{2i} + r_{23} \sum_{i=1}^N x_{1i}x_{3i} + r_{24} \sum_{i=1}^N x_{1i}^2x_2 - \sum_{i=1}^N x_{1i}\dot{x}_{2i} = 0 \\ \frac{\partial J}{\partial r_{22}} &= r_{21} \sum_{i=1}^N x_{1i}x_{2i} + r_{22} \sum_{i=1}^N x_{2i}^2 + r_{23} \sum_{i=1}^N x_{2i}x_{3i} + r_{24} \sum_{i=1}^N x_{1i}x_2^2 - \sum_{i=1}^N x_{2i}\dot{x}_{2i} = 0 \\ \frac{\partial J}{\partial r_{23}} &= r_{21} \sum_{i=1}^N x_{1i}x_{3i} + r_{22} \sum_{i=1}^N x_{2i}x_{3i} + r_{23} \sum_{i=1}^N x_{3i}^2 + r_{24} \sum_{i=1}^N x_{1i}x_2x_{3i} - \sum_{i=1}^N x_{3i}\dot{x}_{2i} = 0 \\ \frac{\partial J}{\partial r_{24}} &= r_{21} \sum_{i=1}^N x_{1i}^2x_{2i} + r_{22} \sum_{i=1}^N x_{1i}x_{2i}^2 + r_{23} \sum_{i=1}^N x_{1i}x_{2i}x_{3i} + r_{24} \sum_{i=1}^N x_{1i}^2x_2^2 - \sum_{i=1}^N x_{1i}x_{2i}\dot{x}_{2i} = 0 \end{aligned}$$

To obtain:

$$\begin{aligned}
\sum_{i=1}^N x_{1i} \dot{x}_{2i} &= r_{21} \sum_{i=1}^N x_{1i}^2 + r_{22} \sum_{i=1}^N x_{1i} x_{2i} + r_{23} \sum_{i=1}^N x_{1i} x_{3i} + r_{24} \sum_{i=1}^N x_{1i}^2 x_2 \\
\sum_{i=1}^N x_{2i} \dot{x}_{2i} &= r_{21} \sum_{i=1}^N x_{1i} x_{2i} + r_{22} \sum_{i=1}^N x_{2i}^2 + r_{23} \sum_{i=1}^N x_{2i} x_{3i} + r_{24} \sum_{i=1}^N x_{1i} x_2^2 \\
\sum_{i=1}^N x_{3i} \dot{x}_{2i} &= r_{21} \sum_{i=1}^N x_{1i} x_{3i} + r_{22} \sum_{i=1}^N x_{2i} x_{3i} + r_{23} \sum_{i=1}^N x_{3i}^2 + r_{24} \sum_{i=1}^N x_{1i} x_2 x_{3i} \\
\sum_{i=1}^N x_{1i} x_{2i} \dot{x}_{2i} &= r_{21} \sum_{i=1}^N x_{1i}^2 x_{2i} + r_{22} \sum_{i=1}^N x_{1i} x_{2i}^2 + r_{23} \sum_{i=1}^N x_{1i} x_{2i} x_{3i} + r_{24} \sum_{i=1}^N x_{1i}^2 x_2^2
\end{aligned}$$

3.3.2 Second matrices set up

$$\begin{pmatrix} \sum_{i=1}^N x_{1i}^2 & \sum_{i=1}^N x_{1i} x_{2i} & \sum_{i=1}^N x_{1i} x_{3i} & \sum_{i=1}^N x_{1i}^2 x_2 \\ \sum_{i=1}^N x_{1i} x_{2i} & \sum_{i=1}^N x_{2i}^2 & \sum_{i=1}^N x_{2i} x_{3i} & \sum_{i=1}^N x_{1i} x_2^2 \\ \sum_{i=1}^N x_{1i} x_{3i} & \sum_{i=1}^N x_{2i} x_{3i} & \sum_{i=1}^N x_{3i}^2 & \sum_{i=1}^N x_{1i} x_2 x_{3i} \\ \sum_{i=1}^N x_{1i}^2 x_{2i} & \sum_{i=1}^N x_{1i} x_{2i}^2 & \sum_{i=1}^N x_{1i} x_{2i} x_{3i} & \sum_{i=1}^N x_{1i}^2 x_2^2 \end{pmatrix} \begin{pmatrix} r_{21} \\ r_{22} \\ r_{23} \\ r_{24} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N x_{1i} \dot{x}_{2i} \\ \sum_{i=1}^N x_{2i} \dot{x}_{2i} \\ \sum_{i=1}^N x_{3i} \dot{x}_{2i} \\ \sum_{i=1}^N x_{1i} x_{2i} \dot{x}_{2i} \end{pmatrix}.$$

3.4 Third set

3.4.1 Third Least-Squares Equation Work

For the third performance index we get:

$$J = \frac{1}{2} \sum_{i=1}^N [r_{31} x_{2i} + r_{32} x_{4i} x_{5i} + r_{33} x_{3i}^2 - \dot{x}_{3i}]^2$$

After setting the derivatives with respect to their respective parameters to zero:

$$\begin{aligned}
\frac{\partial J}{\partial r_{31}} &= r_{31} \sum_{i=1}^N x_{2i}^2 + r_{32} \sum_{i=1}^N x_{2i} x_{4i} x_{5i} + r_{33} \sum_{i=1}^N x_{2i} x_{3i}^2 - \sum_{i=1}^N x_{2i} \dot{x}_{3i} = 0 \\
\frac{\partial J}{\partial r_{32}} &= r_{31} \sum_{i=1}^N x_{2i} x_{4i} x_{5i} + r_{32} \sum_{i=1}^N x_{4i}^2 x_{5i}^2 + r_{33} \sum_{i=1}^N x_{3i}^2 x_{4i} x_{5i} - \sum_{i=1}^N x_{4i} x_{5i} \dot{x}_{3i} = 0 \\
\frac{\partial J}{\partial r_{33}} &= r_{31} \sum_{i=1}^N x_{2i} x_{3i}^2 + r_{32} \sum_{i=1}^N x_{3i}^2 x_{4i} x_{5i} + r_{33} \sum_{i=1}^N x_{3i}^4 - \sum_{i=1}^N x_{3i}^2 \dot{x}_{3i} = 0
\end{aligned}$$

We obtain:

$$\begin{aligned}
\sum_{i=1}^N x_{2i} \dot{x}_{3i} &= r_{31} \sum_{i=1}^N x_{2i}^2 + r_{32} \sum_{i=1}^N x_{2i} x_{4i} x_{5i} + r_{33} \sum_{i=1}^N x_{2i} x_{3i}^2 \\
\sum_{i=1}^N x_{4i} x_{5i} \dot{x}_{3i} &= r_{31} \sum_{i=1}^N x_{2i} x_{4i} x_{5i} + r_{32} \sum_{i=1}^N x_{4i}^2 x_{5i}^2 + r_{33} \sum_{i=1}^N x_{4i} x_{5i} x_{3i}^2 \\
\sum_{i=1}^N x_{3i}^2 \dot{x}_{3i} &= r_{31} \sum_{i=1}^N x_{2i} x_{3i}^2 + r_{32} \sum_{i=1}^N x_{3i}^2 x_{4i} x_{5i} + r_{33} \sum_{i=1}^N x_{3i}^4
\end{aligned}$$

3.4.2 Third matrices set up

$$\begin{pmatrix} \sum_{i=1}^N x_{2i}^2 & \sum_{i=1}^N x_{2i} x_{4i} x_{5i} & \sum_{i=1}^N x_{3i}^2 \\ \sum_{i=1}^N x_{2i} x_{4i} x_{5i} & \sum_{i=1}^N x_{4i}^2 x_{5i}^2 & \sum_{i=1}^N x_{3i}^2 x_{4i} x_{5i} \\ \sum_{i=1}^N x_{2i} x_{3i}^2 & \sum_{i=1}^N x_{3i}^2 x_{4i} x_{5i} & \sum_{i=1}^N x_{3i}^4 \end{pmatrix} \begin{pmatrix} r_{31} \\ r_{32} \\ r_{33} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N x_{2i} \dot{x}_{3i} \\ \sum_{i=1}^N x_{4i} x_{5i} \dot{x}_{3i} \\ \sum_{i=1}^N x_{3i}^2 \dot{x}_{3i} \end{pmatrix}.$$

3.5 Fourth set

3.5.1 Fourth Least-Squares Equation Work

For the fourth performance index we get:

$$J = \frac{1}{2} \sum_{i=1}^N [r_{41}x_1 + r_{42}x_{3i} + r_{43}x_{4i}x_{5i} + r_{44}x_{4i}^2 - \dot{x}_{4i}]^2$$

After setting the derivatives with respect to their respective parameters to zero:

$$\begin{aligned} \frac{\partial J}{\partial r_{41}} &= r_{41} \sum_{i=1}^N x_1^2 + r_{42} \sum_{i=1}^N x_{1i}x_{3i} + r_{43} \sum_{i=1}^N x_{1i}x_{4i}x_{5i} + r_{44} \sum_{i=1}^N x_{1i}x_{4i}^2 - \sum_{i=1}^N x_{1i}\dot{x}_{4i} = 0 \\ \frac{\partial J}{\partial r_{42}} &= r_{41} \sum_{i=1}^N x_{1i}x_{3i} + r_{42} \sum_{i=1}^N x_{3i}^2 + r_{43} \sum_{i=1}^N x_{3i}x_{4i}x_{5i} + r_{44} \sum_{i=1}^N x_{3i}x_{4i}^2 - \sum_{i=1}^N x_{3i}\dot{x}_{4i} = 0 \\ \frac{\partial J}{\partial r_{43}} &= r_{41} \sum_{i=1}^N x_{1i}x_{4i}x_{5i} + r_{42} \sum_{i=1}^N x_{3i}x_{4i}x_{5i} + r_{43} \sum_{i=1}^N x_{4i}^2x_{5i}^2 + r_{44} \sum_{i=1}^N x_{4i}^3x_{5i} - \sum_{i=1}^N x_{4i}x_{5i}\dot{x}_{4i} = 0 \\ \frac{\partial J}{\partial r_{44}} &= r_{41} \sum_{i=1}^N x_{1i}x_{4i}^2 + r_{42} \sum_{i=1}^N x_{3i}x_{4i}^2 + r_{43} \sum_{i=1}^N x_{4i}^3x_{5i} + r_{44} \sum_{i=1}^N x_{4i}^4 - \sum_{i=1}^N x_{4i}^2\dot{x}_{4i} = 0 \end{aligned}$$

We obtain:

$$\begin{aligned} \sum_{i=1}^N x_{1i}\dot{x}_{4i} &= r_{41} \sum_{i=1}^N x_1^2 + r_{42} \sum_{i=1}^N x_{1i}x_{3i} + r_{43} \sum_{i=1}^N x_{1i}x_{4i}x_{5i} + r_{44} \sum_{i=1}^N x_{1i}x_{4i}^2 \\ \sum_{i=1}^N x_{3i}\dot{x}_{4i} &= r_{41} \sum_{i=1}^N x_{1i}x_{3i} + r_{42} \sum_{i=1}^N x_{3i}^2 + r_{43} \sum_{i=1}^N x_{3i}x_{4i}x_{5i} + r_{44} \sum_{i=1}^N x_{3i}x_{4i}^2 \\ \sum_{i=1}^N x_{4i}x_{5i}\dot{x}_{4i} &= r_{41} \sum_{i=1}^N x_{1i}x_{4i}x_{5i} + r_{42} \sum_{i=1}^N x_{3i}x_{4i}x_{5i} + r_{43} \sum_{i=1}^N x_{4i}^2x_{5i}^2 + r_{44} \sum_{i=1}^N x_{4i}^3x_{5i} \\ \sum_{i=1}^N x_{4i}^2\dot{x}_{4i} &= r_{41} \sum_{i=1}^N x_{1i}x_{4i}^2 + r_{42} \sum_{i=1}^N x_{3i}x_{4i}^2 + r_{43} \sum_{i=1}^N x_{4i}^3x_{5i} + r_{44} \sum_{i=1}^N x_{4i}^4 \end{aligned}$$

3.5.2 Fourth matrices set up

$$\begin{pmatrix} \sum_{i=1}^N x_1^2 & \sum_{i=1}^N x_{1i}x_{3i} & \sum_{i=1}^N x_{1i}x_{4i}x_{5i} & \sum_{i=1}^N x_{1i}x_{4i}^2 \\ \sum_{i=1}^N x_{1i}x_{3i} & \sum_{i=1}^N x_{3i}^2 & \sum_{i=1}^N x_{3i}x_{4i}x_{5i} & \sum_{i=1}^N x_{3i}x_{4i}^2 \\ \sum_{i=1}^N x_{1i}x_{4i}x_{5i} & \sum_{i=1}^N x_{3i}x_{4i}x_{5i} & \sum_{i=1}^N x_{4i}^2x_{5i}^2 & \sum_{i=1}^N x_{4i}^3x_{5i} \\ \sum_{i=1}^N x_{1i}x_{4i}^2 & \sum_{i=1}^N x_{3i}x_{4i}^2 & \sum_{i=1}^N x_{4i}^3x_{5i} & \sum_{i=1}^N x_{4i}^4 \end{pmatrix} \begin{pmatrix} r_{41} \\ r_{42} \\ r_{43} \\ r_{44} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N x_{1i}\dot{x}_{4i} \\ \sum_{i=1}^N x_{3i}\dot{x}_{4i} \\ \sum_{i=1}^N x_{4i}x_{5i}\dot{x}_{4i} \\ \sum_{i=1}^N x_{4i}^2\dot{x}_{4i} \end{pmatrix}.$$

3.6 Fifth set

3.6.1 Fifth Least-Squares Equation Work

For the fifth performance index we get:

$$J = \frac{1}{2} \sum_{i=1}^N [r_{51}x_{2i} + r_{52}x_{3i} + r_{53}x_{4i}x_{5i} + r_{54}x_{5i} - \dot{x}_{5i}]^2$$

After setting the derivatives with respect to their respective parameters to zero:

$$\begin{aligned} \frac{\partial J}{\partial r_{51}} &= r_{51} \sum_{i=1}^N x_{2i}^2 + r_{52} \sum_{i=1}^N x_{2i}x_{3i} + r_{53} \sum_{i=1}^N x_{2i}x_{4i}x_{5i} + r_{54} \sum_{i=1}^N x_{2i}x_{5i} - \sum_{i=1}^N x_{2i}x_{3i}\dot{x}_{5i} = 0 \\ \frac{\partial J}{\partial r_{52}} &= r_{51} \sum_{i=1}^N x_{2i}x_{3i} + r_{52} \sum_{i=1}^N x_{3i}^2 + r_{53} \sum_{i=1}^N x_{3i}x_{4i}x_{5i} + r_{54} \sum_{i=1}^N x_{3i}x_{5i} - \sum_{i=1}^N x_{3i}\dot{x}_{5i} = 0 \\ \frac{\partial J}{\partial r_{53}} &= r_{51} \sum_{i=1}^N x_{2i}x_{4i} + r_{52} \sum_{i=1}^N x_{3i}x_{4i} + r_{53} \sum_{i=1}^N x_{4i}^2x_{5i}^2 + r_{54} \sum_{i=1}^N x_{4i}x_{5i}^2 - \sum_{i=1}^N x_{4i}x_{5i}\dot{x}_{5i} = 0 \\ \frac{\partial J}{\partial r_{54}} &= r_{51} \sum_{i=1}^N x_{2i}x_{5i} + r_{52} \sum_{i=1}^N x_{3i}x_{5i} + r_{53} \sum_{i=1}^N x_{4i}x_{5i}^2 + r_{54} \sum_{i=1}^N x_{5i}^2 - \sum_{i=1}^N x_{5i}\dot{x}_{5i} = 0 \end{aligned}$$

We obtain:

$$\begin{aligned} \sum_{i=1}^N x_{2i}\dot{x}_{5i} &= r_{51} \sum_{i=1}^N x_{2i}^2 + r_{52} \sum_{i=1}^N x_{2i}x_{3i} + r_{53} \sum_{i=1}^N x_{2i}x_{4i} + r_{54} \sum_{i=1}^N x_{2i}x_{5i} \\ \sum_{i=1}^N x_{3i}\dot{x}_{5i} &= r_{51} \sum_{i=1}^N x_{2i}x_{3i} + r_{52} \sum_{i=1}^N x_{3i}^2 + r_{53} \sum_{i=1}^N x_{3i}x_{4i}x_{5i} + r_{54} \sum_{i=1}^N x_{3i}x_{5i} \\ \sum_{i=1}^N x_{4i}x_{5i}\dot{x}_{5i} &= r_{51} \sum_{i=1}^N x_{2i}x_{4i}x_{5i} + r_{52} \sum_{i=1}^N x_{3i}x_{4i}x_{5i} + r_{53} \sum_{i=1}^N x_{4i}^2x_{5i}^2 + r_{54} \sum_{i=1}^N x_{4i}x_{5i}^2 \\ \sum_{i=1}^N x_{5i}\dot{x}_{5i} &= r_{51} \sum_{i=1}^N x_{2i}x_{5i} + r_{52} \sum_{i=1}^N x_{3i}x_{5i} + r_{53} \sum_{i=1}^N x_{4i}x_{5i}^2 + r_{54} \sum_{i=1}^N x_{5i}^2 \end{aligned}$$

3.6.2 Fifth matrices set up

$$\begin{pmatrix} \sum_{i=1}^N x_{2i}^2 & \sum_{i=1}^N x_{2i}x_{3i} & \sum_{i=1}^N x_{2i}x_{4i} + r_{54} \sum_{i=1}^N x_{2i}x_{5i} & \sum_{i=1}^N x_{3i}x_{5i} \\ \sum_{i=1}^N x_{2i}x_{3i} & \sum_{i=1}^N x_{3i}^2 & \sum_{i=1}^N x_{3i}x_{4i}x_{5i} & \sum_{i=1}^N x_{4i}x_{5i}^2 \\ \sum_{i=1}^N x_{2i}x_{4i}x_{5i} & \sum_{i=1}^N x_{3i}x_{4i}x_{5i} & \sum_{i=1}^N x_{4i}^2x_{5i}^2 & \sum_{i=1}^N x_{4i}x_{5i}^2 \\ \sum_{i=1}^N x_{2i}x_{5i} & \sum_{i=1}^N x_{3i}x_{5i} & \sum_{i=1}^N x_{4i}x_{5i}^2 & \sum_{i=1}^N x_{5i}^2 \end{pmatrix} \begin{pmatrix} r_{51} \\ r_{52} \\ r_{53} \\ r_{54} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N x_{2i}\dot{x}_{5i} \\ \sum_{i=1}^N x_{3i}\dot{x}_{5i} \\ \sum_{i=1}^N x_{4i}x_{5i}\dot{x}_{5i} \\ \sum_{i=1}^N x_{5i}\dot{x}_{5i} \end{pmatrix}.$$

3.7 Gauss-Seidel Method

One of the methods that was used to find the parameters is the Gauss-Seidel Method. This method is used when conventional least squares methods cannot be used, such as when one has an ill-conditioned matrix. In order to use this method one still needs a matrix that is derived using the derivatives of the least squares equation at hand with respect to each of the parameters. Once one has the matrix we can set up the equation:

$$[A][x] = [b]$$

where in this case $[x]$ is the parameters, r_{ij} we are searching for and $[b]$ are the sums of the data, $\sum_{i=1}^N x_i$ from the respective variables used.

To use Gauss-Seidel Method we shall solve equation 3.1 iteratively assuming the initial x -values are zero. Once we find the new x -values we then plug these new values in to find a new set of x -values. We do this iteratively until we hit a stopping criteria, ϵ , which is the error between the newest value of x_i and the last value of x_i .

To illustrate this method we shall use a 3×3 symmetric matrix for $[A]$ and a vector for $[b]$ in the equation:

$$\begin{pmatrix} 3 & 6 & 5 \\ 6 & 2 & 7 \\ 5 & 7 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 6 \end{pmatrix}.$$

Now that we have this set up we can solve for each parameter:

$$\begin{aligned}
 x_1^k &= \frac{5 - 6x_2^{k-1} - 5x_3^{k-1}}{3} \\
 x_2^k &= \frac{3 - 6x_1^k - 7x_3^{k-1}}{2} \\
 x_3^k &= \frac{6 - 5x_1^k - 7x_2^k}{1}
 \end{aligned}$$

where the k marks the iteration version of the parameter.

Now we just use the $[x]_i^k$ equations for the iterations. For the first iteration the $[x]^k$ vector shall be:

$$[x]^0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Now the results for the first iteration:

$$\begin{aligned}
 x_1^1 &= \frac{5 - 6 * 0 - 5 * 0}{3} = \frac{5}{3} \\
 x_2^1 &= \frac{3 - 6 * \frac{5}{3} - 7 * 0}{2} = \frac{-7}{2} \\
 x_3^1 &= \frac{6 - 5 * \frac{5}{3} - 7 * \frac{-7}{2}}{1} = \frac{133}{6}
 \end{aligned}$$

To get the second iteration we just have to do the same thing but using the new vector, $[x]^1$, instead of $[x]^0$ then:

$$[x]^2 = \begin{pmatrix} \frac{-509}{6} \\ \frac{2141}{6} \\ \frac{-6203}{3} \end{pmatrix}$$

Thus you have your second iteration of values for $[x]$ using Gauss-Seidel Method.

3.8 Results

This part of the research resulted in the network dynamics for this system.

3.8.1 Plots of the evolution of the network

These plots show data of each variable and the rate of change for each variable.

In Figure 3.1 we see the data gathered for the x_1 variable using the programs described above. The first plot shows the actual data generated using the given constraints described above. The following four plots show the results of the data generated of the other four variables.

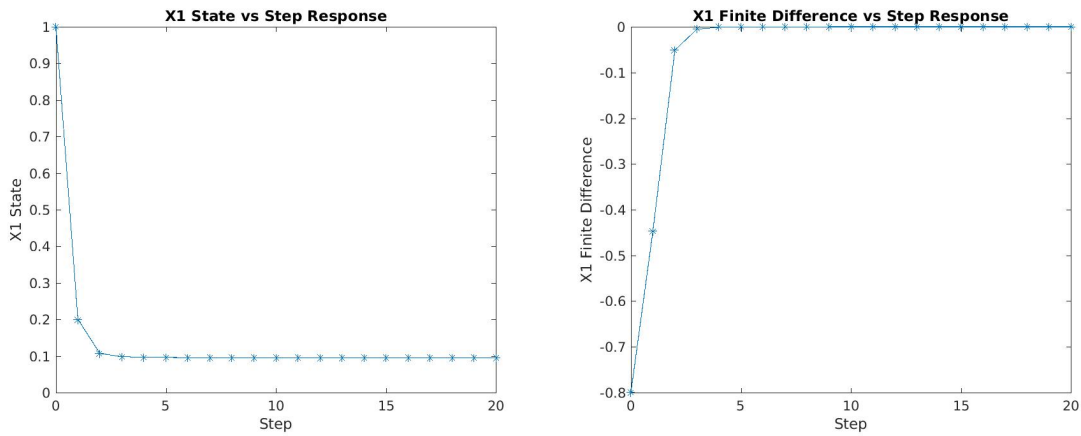


Figure 3.1: State of X_1 and its finite difference.

3.8.2 The comparison plots that show the observed dynamics with the calculated dynamics of the system

The following plots illustrate our results, which come after generating data and a parameter search involving least squares methods. The plots show the contrast between the data generated

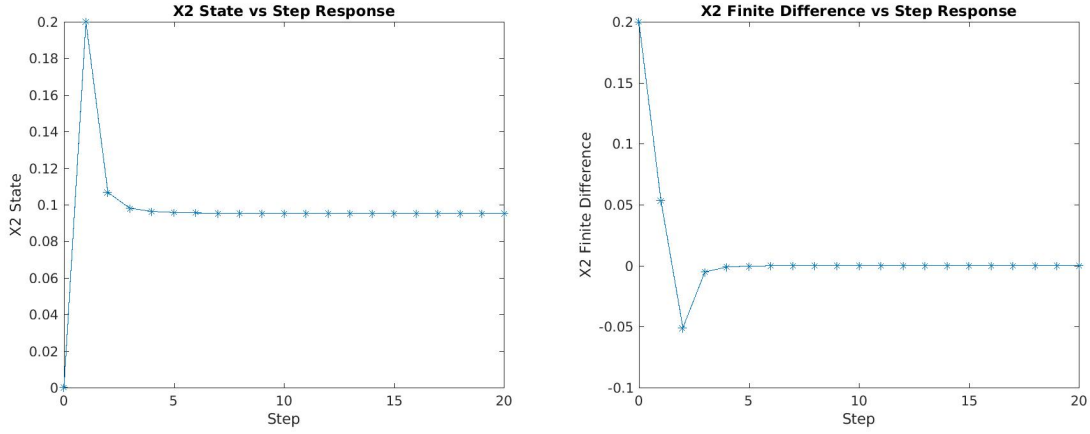


Figure 3.2: State of X_2 and its finite difference.

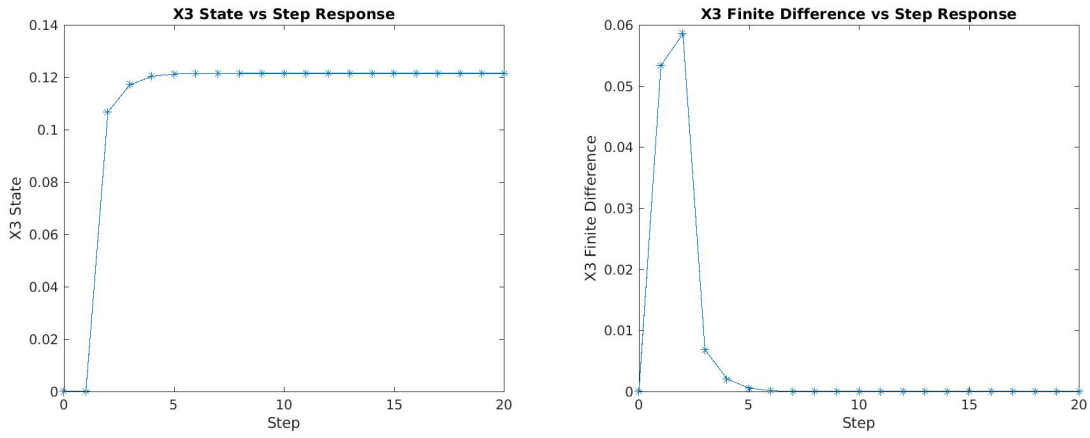


Figure 3.3: State of X_3 and its finite difference.

versus the approximated differential equations using the newly found parameters.

3.8.3 The System of Differential Equations Found

$$\frac{\partial x_1}{\partial t} = -0.6 * x_1 + 0.1 * x_2 + 0.393 * x_5$$

$$\frac{\partial x_2}{\partial t} = 0.5 * x_1 - 0.6 * x_2 + .2 * x_3 - 1.8 * x_1 * x_2$$

$$\frac{\partial x_3}{\partial t} = 0.2 * x_2 + 0.1 * x_4 * x_5 - 1.395 * x_3^2$$

$$\frac{\partial x_4}{\partial t} = 0.3 * x_1 + 0.1 * x_3 - 0.5 * x_4 * x_5 - 2.139 * x_4^2$$

$$\frac{\partial x_5}{\partial t} = 0.3 * x_2 + 0.1 * x_3 - 0.8 * x_4 * x_5 - 0.236 * x_5$$

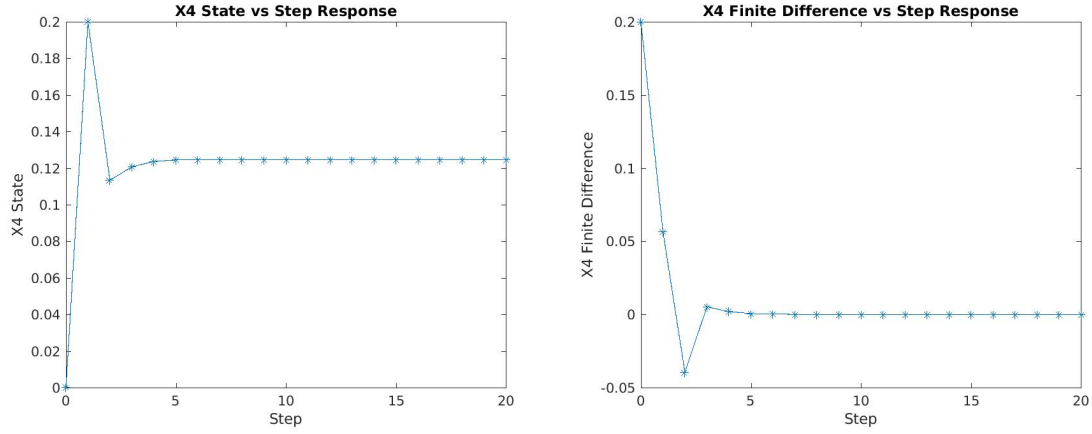


Figure 3.4: State of X_4 and its finite difference.

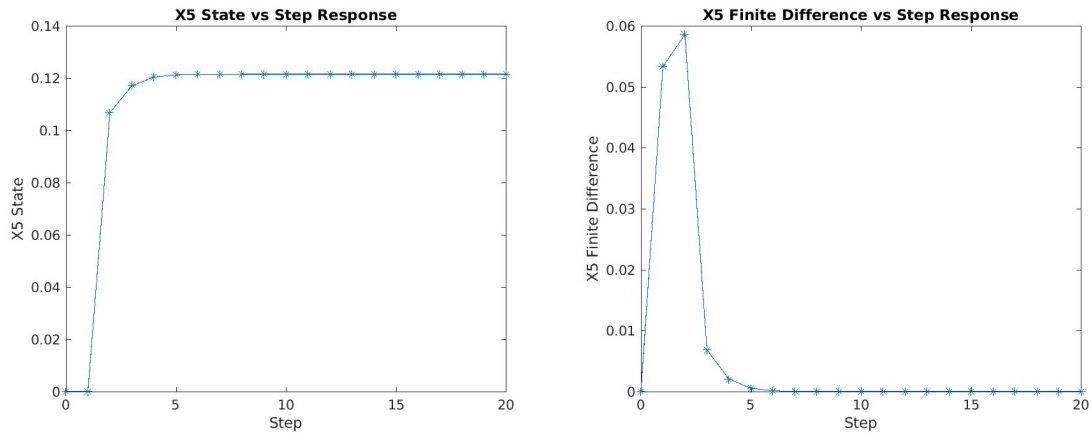


Figure 3.5: State of X_5 and its finite difference.

The following figures show the plots of the systems of equations found using Runge-Kutta methods overlayed with the original collected data.

3.8.4 Stability Analysis

Now that we have approximated the system of differential equations that define the network an analysis of the stability is performed. To do this we have to take the Jacobian of the system by

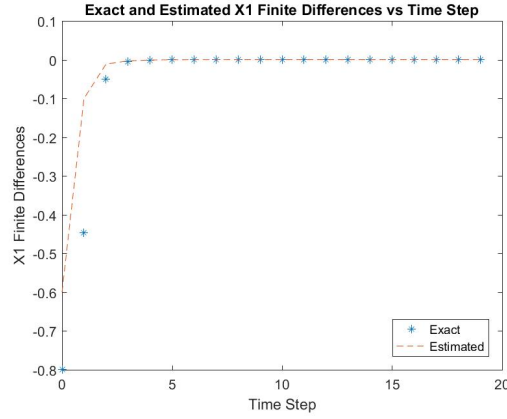


Figure 3.6: Actual finite differences for X_1 and the finite differences estimated from the parameters found using least squares methods.

first setting the system of differential equations to zero:

$$-0.6 * x_1 + 0.1 * x_2 + 0.393 * x_5 = 0$$

$$0.5 * x_1 - 0.6 * x_2 + 0.2 * x_3 - 1.8 * x_1 * x_2 = 0$$

$$0.2 * x_2 + 0.1 * x_4 * x_5 - 1.395 * x_3^2 = 0$$

$$0.3 * x_1 + 0.1 * x_3 - 0.5 * x_4 * x_5 - 2.139 * x_4^2 = 0$$

$$0.3 * x_2 + 0.1 * x_3 - 0.8 * x_4 * x_5 - 0.236 * x_5 = 0$$

Now to construct the Jacobian matrix we take the first derivative of each equation with respect to the five variables that make up the network to create:

$$\begin{bmatrix} -0.6 & 0.1 & 0 & 0 & 0.393 \\ 0.5 - 1.8 * x_2 & -0.6 - 1.8 * x_1 & 0.2 & 0 & 0 \\ 0 & 0.2 & -2.790 * x_3 & 0.1 * x_5 & 0.1 * x_4 \\ 0.3 & 0 & 0.1 & -0.5 * x_5 - 4.278 * x_4 & -0.5 * x_4 \\ 0 & 0.3 & 0.1 & -0.8 * x_5 & -0.8 * x_4 - 0.236 \end{bmatrix}$$

After constructing the Jacobian we then find the steady-state approximation for each variable using the last data points from the variable state versus step response plots, which was the stationary

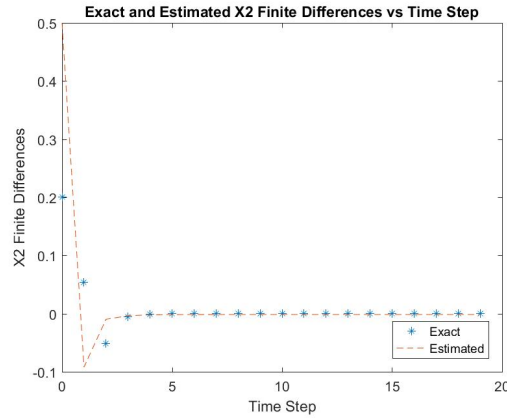


Figure 3.7: Actual finite differences for X_2 and the finite differences estimated from the parameters found using least squares methods.

distribution:

$$X = \begin{bmatrix} 0.09199491315 & 0.09109270746 & 0.1187114309 & 0.1228228975 & 0.1172714431 \end{bmatrix}$$

. This steady-state approximation will be considered the equilibrium point of the system of differential equations. We then find the eigenvectors and their corresponding eigenvalues of the Jacobian using the equilibrium point approximations using matlab to get the following eigen-pairs (eigenvector and eigenvalue):

$$\begin{bmatrix} -0.0722 - 0.3217i \\ -0.7276 + 0.0000i \\ 0.2816 + 0.0645i \\ -0.1449 + 0.3553i \\ 0.3307 + 0.1476i \\ -0.5205 + 0.0000i \\ -0.3877 + 0.0000i \\ -0.3978 + 0.0000i \\ -0.3413 + 0.0000i \\ -0.5513 + 0.0000i \end{bmatrix} ; -0.8097 + 0.1308i \quad \begin{bmatrix} -0.0722 + 0.3217i \\ -0.7276 + 0.0000i \\ 0.2816 - 0.0645i \\ -0.1449 - 0.3553i \\ 0.3307 - 0.1476i \\ 0.4258 + 0.2925i \\ 0.2521 + 0.0537i \\ -0.3277 - 0.3177i \\ 0.6510 + 0.0000i \\ 0.0541 + 0.1779i \end{bmatrix} ; -0.8097 - 0.1308i$$

$$\begin{bmatrix} -0.0722 - 0.3217i \\ -0.7276 + 0.0000i \\ 0.2816 + 0.0645i \\ -0.1449 + 0.3553i \\ 0.3307 + 0.1476i \\ -0.5205 + 0.0000i \\ -0.3877 + 0.0000i \\ -0.3978 + 0.0000i \\ -0.3413 + 0.0000i \\ -0.5513 + 0.0000i \end{bmatrix} ; -0.1092 + 0.0000i \quad \begin{bmatrix} -0.0722 + 0.3217i \\ -0.7276 + 0.0000i \\ 0.2816 - 0.0645i \\ -0.1449 - 0.3553i \\ 0.3307 - 0.1476i \\ 0.4258 + 0.2925i \\ 0.2521 + 0.0537i \\ -0.3277 - 0.3177i \\ 0.6510 + 0.0000i \\ 0.0541 + 0.1779i \end{bmatrix} ; -0.4433 + 0.0692i$$

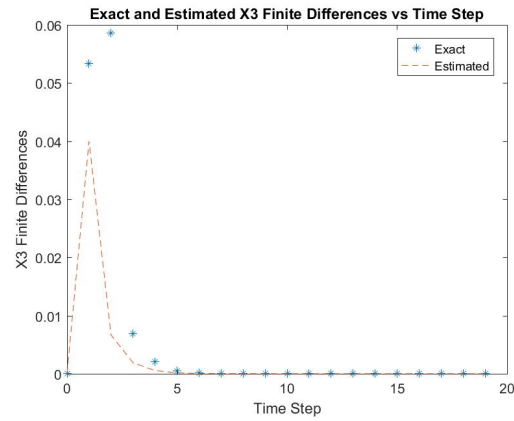


Figure 3.8: Actual finite differences for X_3 and the finite differences estimated from the parameters found using least squares methods

$$\begin{bmatrix} -0.2925i \\ 0.2521 - 0.0537i \\ -0.3277 + 0.3177i \\ 0.6510 + 0.0000i \\ 0.0541 - 0.1779i \end{bmatrix} ; -0.4433 - 0.0692i.$$

According to a theorem on the linear stability at a point, since the eigenvalues of each eigenpair have negative valued real parts the system of differential equations is stable.

What having a stable equilibrium point means for this system of differential equations is that the network would not easily be unstabilized given disturbances from outside the network.

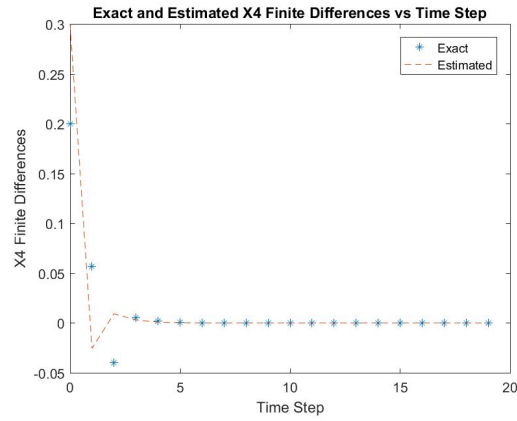


Figure 3.9: Actual finite differences for X_4 and the finite differences estimated from the parameters found using least squares methods

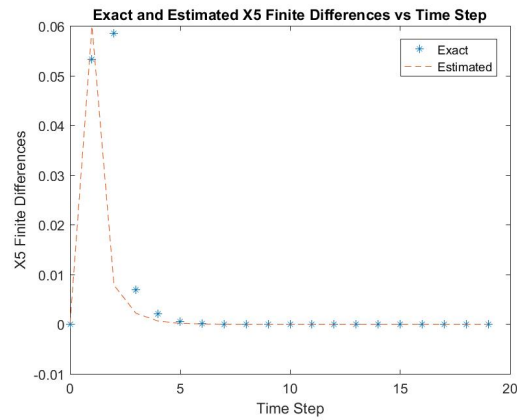


Figure 3.10: Actual finite differences for X_5 and the finite differences estimated from the parameters found using least squares methods

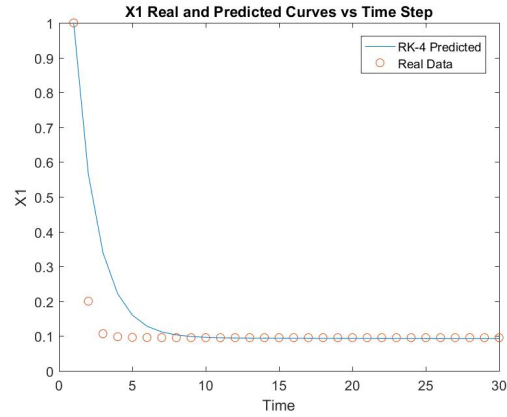


Figure 3.11: Actual state of X_1 and the predicted state found using the RK-4 method.

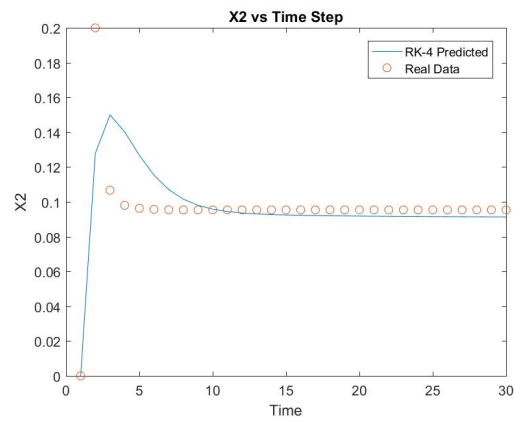


Figure 3.12: Actual state of X_2 and the predicted state found using the RK-4 method.

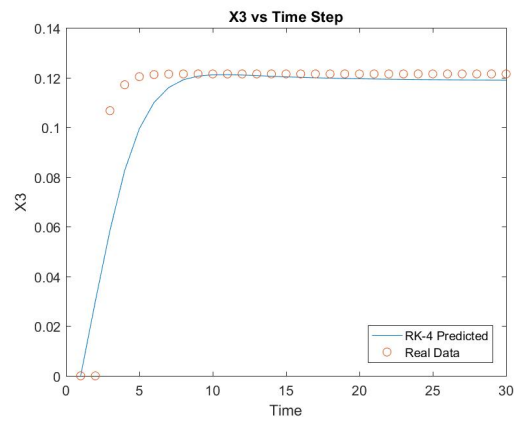


Figure 3.13: Actual state of X_3 and the predicted state found using the RK-4 method.

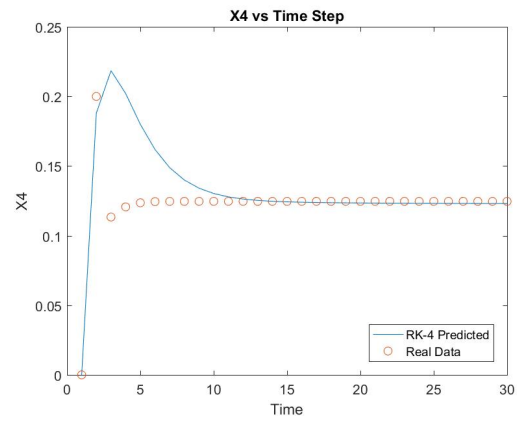


Figure 3.14: Actual state of X_4 and the predicted state found using the RK-4 method.

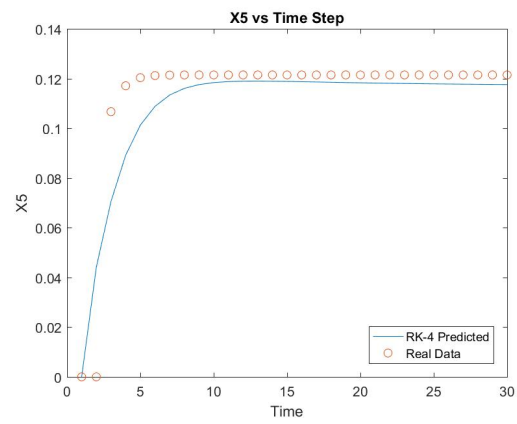


Figure 3.15: Actual state of X_5 and the predicted state found using the RK-4 method.

CHAPTER 4

IN CLOSING

The resulting system of differential equations and stability analysis of the system of differential equations implies that this gene network is stable. As viewed in a biological sense this means that this gene network should be very hard to disrupt so that the cell cycle does not have problems on a regular basis. That being said there are a lot of different mechanisms within the cell, outside this gene network, which help in maintaining the network's stability.

Future work can be done to find the breaking points for this gene network and other networks like it. As was stated earlier, this research is meant to help find a treatment schedule for breast cancer patients. The gathered data serves as a starting point for finding such schedules.

REFERENCES

- [1] Marcos Malumbres, Mariano Barbacid *Cell cycle, CDKs and cancer: a changing paradigm*, 9th Vol., Nature. (2009). www.nature.com/reviews/cancer
- [2] Xiang H.-F. Zhang, Mario Giuliano, Meghana V. Trivedi, Rachel Schiff, and C. Kent Osborne *Metastasis Dormancy in Estrogen ReceptorPositive Breast Cancer*, clincancerres.aacrjournals.org on April 21, 2016.
- [3] C. Kent Osborne, M.D. *Tamoxifen in the Treatment of Breast Cancer*, Volume 339, Number 22. The New England Journal of Medicine. Massachusetts Medical Society. 1998.
- [4] Xiangwei Wu and Scott M. Lippman *An intermittent approach for cancer chemoprevention*, Volume 11, December 2011. Nature. (2011). www.nature.com/reviews/cancer
- [5] Kevin Leder, Ken Pitter, Quincey LaPlant, Dolores Hambarzumyan, Brian D. Ross, Timothy A. Chan, Eric C. Holland, and Franziska Michor *Mathematical Modeling of PDGF-Driven Glioblastoma Reveals Optimized Radiation Dosing Schedules*, <http://dx.doi.org/10.1016/j.cell.2013.12.029>
- [6] George W. Sledge, Eleftherios P. Mamounas, Gabriel N. Hortobagyi, Harold J. Burstein, Pamela J. Goodwin, Lunenfeld-Tanenbaum, and Antonio C. Wolff *Past, Present, and Future Challenges in Breast Cancer Treatment*, Volume 32, Number 19 Journal of Clinical Oncology July 1, 2014.