STUDY ON DYNAMIC MODELLING OF METABOLIC FLUX ANALYSIS AND APPLICATION IN METABOLIC-RELATED DISEASE MODEL

HUO YUNZHANG

PhD

The Hong Kong Polytechnic University

2020

The Hong Kong Polytechnic University

Department of Industrial and Systems Engineering

Study on Dynamic Modelling of Metabolic Flux Analysis and Application in Metabolic-related Disease Model

HUO Yunzhang

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Aug 2019

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(Signed)

HUO Yunzhang_____(Name of student)

Abstract

In recent years, the concepts and theories of systems biology have developed rapidly. Systems biology is an approach to integrating the knowledge of each level and each component. Among these levels, fluxome is an increasingly important component in addition to well-known genomes, transcriptomes, proteomes, and metabolomes. The study of this system has produced a new omics approach, metabolic fluxomics. The focus is on determining of the direction and value of the metabolic flux within the target system. The information is important for characterizing the physiological characteristics of biological systems and assessing genetic and environmental impacts on cells. The main research tool of metabolic fluxomics is metabolic flux analysis (MFA). MFA, a key technology in bioinformatics, is an effective way of analyzing the entire metabolic system by measuring fluxes. Many existing MFA approaches are based on differential equations, which are complicated to solve mathematically. Therefore, MFA requires some simple approaches to further investigate metabolism.

In this thesis, a Continuous-time Markov chain approach was proposed to MFA, called MMFA approach, and transformed the MFA problem into a set of quadratic equations by analyzing the transition probability of each carbon atom in the entire metabolic system. Unlike other methods, MMFA analyzes the metabolic model only through the transition probability. This approach is generic and could be applied to any metabolic system if all the reaction mechanisms in the system are known. The MMFA approach was applied to the pentose phosphate pathway and

the results were compared to several chemical reaction equilibrium constants from early experiments. In order to apply the approach to larger and more complex central metabolic processes, such as pentose phosphate pathway, glycolysis and tricarboxylic acid (TCA) cycle, the Monte Carlo method was used to solve the problem that not all reaction rate constants *k* can be available. The results indicated that continuous-time Markov chain metabolic flux analysis can be applied to more biological systems using the Monte Carlo method.

Furthermore, this research continued to apply MFA to biological models related to metabolic diseases. The insulin signaling pathway was selected as the model. In regulated cells, insulin receptors are on the cell surface. The receptors, in combination with insulin, lead to a series of changes in the cells (signal transduction pathway). One of the reasons that lead to Type_2 diabetes mellitus is insulin resistance. The drugs do not have a guaranteed effect because the genes themselves have several steady states, while the insulin can work in some steady states but not in all of them. When insulin is injected, once the steady states reached by the genes are similar to those resisted by insulin, it stops functioning. The steady states here refer to the stable steady point with the minimum potential energy in a single cell, which may be different from the definition in biological theory. Multiple homeostasis problems are taken into consideration in the dynamic process, while multiple homeostasis in the insulin receptor is proposed to reveal the cause of insulin resistance in Type 2 diabetes through the reaction dynamic model.

In addition, meta-analysis was applied to gene expression during heart

regeneration in Danio rerio. Danio rerio (zebrafish) is a well-known model vertebrate owing to its ability to regenerate its organs, and this dissertation aimed to generate a global profile of genes and transcripts participating in zebrafish heart regeneration. The discovery of its heart regeneration ability and the preliminary revelation of the underlying molecular and cellular mechanisms are considered a breakthrough in the field of organ regeneration in recent years. However, the molecular mechanisms underlying heart regeneration at the transcript level are unclear. Using statistical analysis of several RNA-seq data from the published literature, the expression data of more than 30,000 genes in zebrafish, during its heart regeneration, were obtained. Gene clustering and mathematical statistics were used to analyze the expression data from different aspects. The results showed a more accurate gene differential expression profiles during zebrafish heart regeneration and obtained the corresponding biological processes by go-cluster analysis. These data might aid in guiding future research on heart regeneration of zebrafish. Our analysis, which is comprehensive and reliable, revealed several genes that have not been linked to the heart regeneration process and might be useful for rebuilding cardiac function in zebrafish.

This research contributes an innovative and generic approach to metabolic flux analysis and reveals some meaningful mechanisms in disease models.

Publication arising from this study

Journal papers

1. **Yunzhang Huo,** and Ping Ji. " Continuous-time Markov chain based flux analysis in metabolism." *Journal of Computational Biology*. 21(9): 691-698.

2. **Yunzhang Huo,** Carman Ka Man Lee and Ping Ji. " Continuous-time Markov chain with Monte Carlo based flux analysis in central metabolism." *Journal of Computational Biology*. (Submitted).

3. **Yunzhang Huo,** Carman Ka Man Lee and Ping Ji. " Multiple homeostasis in the insulin receptor reveals the cause of insulin resistance in Type 2 diabetes." *International Journal of Advances in Science Engineering and Technology.* 7(3): 43-49.

4. **Yunzhang Huo,** Carman Ka Man Lee and Ping Ji. " Meta-Analysis of Gene Expression during Heart Regeneration in Danio rerio." *Cellular and Molecular Biology*. (Submitted).

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my chief supervisor Prof. Carman Lee and co-supervisor Prof. Ji Ping, who are knowledgeable professors from the Hong Kong Polytechnic University. Their great patience, valuable guidance, positive encouragement and continuous support are of great help to complete my Ph.D study. Besides, their obliging souls and positive attitudes towards life really inspire me a lot in my personal life.

Besides my supervisor and co-supervisor, I would like to express my grateful thanks to my mates (Dr. NG KAM HUNG, Dr. Liu Chengli, NG, Mr. Dicky, K.L. KEUNG) for their valuable advices and kind help in my research. Besides, my great thanks go to Dr. Liu Da, Mr. Feng Mingyu, Dr. Liu Zhongwan, Dr. Jiang Wenshuai, Dr. Shen Yichi and Mr. Wang Yue for their helpful instructions and great encouragement.

Finally, I wish to express my deeply grateful to my wife, Ms Sun Wei, and my family members for their selfless love, understanding and support throughout the period of my study.

Table of Contents

List of Figures

List of Table

Chapter 1 Introduction

1.1 Background

Metabolic flux analysis (MFA) is a powerful tool for the characterization of the cellular state by quantification of intracellular fluxes in a metabolic network. In order to overcome the limitation of stoichiometric MFA, ^{13}C MFA based on carbon-labeling experiments has been developed. In this thesis, the principle and the method of ^{13}C MFA are introduced, the development of ^{13}C MFA in experiment and data analysis are summarized, and the important roles of MFA in the research on functional genomics are reviewed. The future developments of MFA are proposed.

Metabolic flux is an essential factor in the cell physiology and a very important parameter of the metabolic pathway. Under steady-state conditions, metabolic flux is in general showed as the metabolites production rates. Metabolic flux analysis (MFA) is based on each reaction's measurement in relation to metabolic pathways. It is a way to determine the metabolic reaction network's metabolic flux distribution based on the measured data and it occupies an important position in metabolic engineering.

By calculating different approaches or metabolic flux distributions under different conditions, the study shows the ability of cell metabolism and expose the genetic modification effect on the state of cell metabolism, thereby providing theoretical basis for more reasonable genetic modification.

MFA combines the substrate absorption and product formation rates, the organisms' biosynthesis demand, the stoichiometry of the reaction, the biosynthesis of intermediate metabolites and other data, providing key quantitative information for rebuilding the cell metabolism. MFA's common concerns are with central carbon metabolic pathways (including the glycolysis pathway, tricarboxylic acid cycle, and pentose phosphate pathway), sometimes also including part of the biosynthetic pathways, such as amino acid biosynthesis.

MFA has been successfully applied to a variety of biological optimization solutions, and has increased the productivity of many processes, such as the production of organic acids, vitamins and ethanol. MFA experienced a rapid development in 10 years, in the experiment, and measuring methods as well as in data analysis and assessment. The related technology is mature and becomes a standard to be widely used diagnostic tool in metabolic engineering.

As more and more gene regulatory networks and metabolic networks are being constructed at the whole genome level, the development of analytical tools and methods for metabolic networks is important and imminent. Many researchers adopted various models and approached the study of metabolic flux analysis in a metabolic dynamic system. However, there are still some research areas that needs to be investigated. Improving the simulation model with chemical kinetic models for metabolic flux analysis (MFA) is an important research problem. The metabolic network can be analyzed from various perspectives. Bailey [Bai91] analyzed the metabolic network by altering pathway distributions and rates in 1991.

Prior to 1991, most of the researchers analyzed metabolic process by the static analysis method. When a large number of kinetic parameters and metabolite concentration parameters are unknown in a large-scale metabolic network, the dynamic method is still not conducive for analysis. Based on the kinetic model and kinetic parameters, the simulation model could be generic for the metabolic network systems, as well as other biological system networks.

Metabolic flux analysis (MFA) is an important approach in metabolic engineering, and the basic method in metabolic network analysis. MFA can calculate the intracellular flux according to the metering model of the main intracellular chemical reactions and mass balance model of intracellular metabolites [Sch12]. The basis of MFA is quasi steady-state approximation: in the fastest stage of product formation, the variation in intracellular concentration values of intermediate metabolites should be 0. According to the mass balance model, n equations of reaction rate based on n intermediate metabolites can be obtained. After calculating the flux of an unknown path through the determination of extracellular metabolites, the distribution model of intracellular flux is formulated. Optimal solution of the interference on glucose metabolism can be used to improve the treatment to metabolic deficiency in humans or plants. There could be many reasons for crops to present a metabolism deficiency, such as the lack of an element that is necessary for enzyme synthesis, or a gene mutation. The findings about an optimal solution can provide some direction to the treatment. The crop is induced to turn to the optimal bypath so that it will not only survive but also reach better

growth. Currently, transgenic crops and animals from a popular research area, and they do provide more food resources. With the purpose of increasing the utilization of metabolic, new approaches may improve the genetic engineering strategies. This study attempts to orientate the crops' metabolism towards a desired direction with specific inducible promoters, inhibition or activation of genes and some other methods.

Metabolic disease is one of the rapidly growing high-risk diseases. The most common diseases include diabetes and diabetic cardiopathy. Diabetes is a complex metabolic disease caused by defects in insulin secretion and insulin resistance. Insulin is a classic metabolic regulating hormone and plays an irreplaceable role in the regulation of glucose and lipid metabolism in the body. Diabetes is a metabolic disorder that has an important impact on the cardiovascular system and can induce diabetes-related cardiovascular diseases, including atherosclerosis, coronary heart disease, hypertension, and heart failure ([Jia18], [Mic17], [Eve15]). The cause of death in 70% to 80% people with diabetes is cardiovascular disease. Therefore, as early as 1999, the American Heart Association made it clear that diabetes is a cardiovascular disease. In 2001, the American Cholesterol Education Program listed diabetes as an equal risk for coronary heart disease [Cle01] and regarded cardiovascular disease as a local tissue manifestation of imbalance in the bodys metabolic homeostasis. Diabetes patients have almost all risk factors associated with cardiovascular disease, such as hyperglycemia, dyslipidemia, hypertension, hypercoagulability, hyperinsulinemia, vascular endothelial

dysfunction, and vascular inflammation. These are the main manifestations of metabolic syndrome.

Cardiovascular insulin signaling plays an important role in cardiovascular protection, and cardiovascular insulin resistance is one of the important pathophysiological mechanisms leading to cardiovascular structural and functional changes, which in turn promote cardiovascular disease. Therefore, some of the humoral factors associated with clinical insulin sensitivity may be potential biomarkers of cardiovascular disease [Tou13].

1.2 Problem Statement

Metabolic flux analysis (MFA) analyzes the distribution of balanced metabolic fluxes based on the principle of conservation of matter in metabolic networks. The input data is mainly derived from the measurement of extracellular flux, which is a traditional method. However, intracellular metabolic fluxes, however, difficult to be directly assessed in living organism but have to be realized from measured quantities. Metabolic flux analysis based on carbon labeling experiments (^{13}C MFA) is a more powerful and more recently developed quantitative method. Labeled carbon measurement data provides a large amount of input information for the method and is the most commonly used method for accurate analysis of metabolic flux distribution. Specifically, in MFA, qualitative biological knowledge is mainly the structure of metabolic networks, metabolic flux constraints and the selection of metabolic targets. Some of these contents can be

directly obtained from the literature, while others are not directly available and need adjustments for the improvement of the analysis. MFA can perform modeling analysis from multiple perspectives, for effective analysis of the metabolic process. The traditional MFA method focuses more on the study of the state of steady state and reverses the reaction process with steady state as a result. But the dynamic process deduced by the forward process can make the MFA method more flexible and can create more complex models that can describe biological processes. Therefore, the multi-angle dynamic model can also better simulate the isotope tracing experiments (such as the ^{13}C labeling experiment), in order to reduce the various experimental costs consumed in the disease model research in order to better simulate the disease model. Disease model is a model proposed to descript the relationship among factors and also identify the empirical phenomenon. Metabolic disease is one of the rapidly growing high-risk diseases. The most common diseases include diabetes and diabetic cardiopathy. Multiple approaches can also be applied in metabolic related disease models. Diabetes and cardiovascular disease are all metabolic-related diseases. Therefore, the innovation and development of the MFA method helps further study and analysis of metabolic-related disease models. The subject of this study is improving the simulation model for metabolic flux analysis (MFA) with a chemical kinetic model, a generic dynamic model is proposed. It helps to analyze the metabolic process when the system reaches steady state. The approach will also be applied to other disease models related to metabolism in order to reveal more valuable information

in biological system networks.

1.3 Research Objectives and Significance

Improving the simulation model for metabolic flux analysis (MFA) with a chemical kinetic model, a generic dynamic model is proposed. It helps to analyze the metabolic process when the system reaches steady state. The approach will also be applied to other disease models related to metabolism in order to reveal more valuable information in biological system networks. This study developed a new method to analyze existing data related to the disease without experiment. The objectives of this research are as follows:

(i) To introduce the continuous-time Markov chain to the simulation of the pentose phosphate pathway for metabolic flux analysis in order to indicate the effectiveness and accuracy of the approach.

(ii) To adopt the continuous-time Markov chain MFA approach to central metabolic processes including the pentose phosphate pathway, glycolysis, and tricarboxylic acid (TCA) cycle in order to realize the approach can be used on larger and more complex network models in metabolism.

(iii) To apply the MFA approach to other disease models related to metabolism. Analyze the biological characteristics in disease models when the system reaches steady state.

(iv) To demonstrate how meta-analysis for gene expression in heart regeneration

process of Danio rerio in order to find several genes that have not been linked to the heart regeneration process and might be useful for rebuilding cardiac function in zebrafish.

This research is meaningful and innovative for adopting MFA approaches for formulating disease models. Multi-angle approaches could allow MFA to better describe and analyze metabolic processes as well as metabolic-related disease models. It can also provide more effective and accurate information and data for system biological processes. The application of MFA to metabolic-related disease models can be an effective in solving real-world disease problems and drug development.

1.4 Organization of the Thesis

Seven chapters are included in this thesis. Chapter 1 provides an overall background about metabolic flux analysis (MFA), insulin signaling pathway and heart regeneration. The research objectives are also proposed.

In chapter 2, relevant concepts and different methods regarding metabolic flux analysis are explained. Reviews on metabolic homeostasis and heart regeneration are also provided.

In chapter 3, continuous-time Markov chain to MFA is proposed and applied the approach to the pentose phosphate pathway. The results were compared with several chemical reaction equilibrium constants from early experiments. Different reactants and products in the pentose phosphate pathway are related to other

metabolic systems. In addition, the relative concentration was evaluated with the reaction equilibrium constant via different simulation scenarios in order to demonstrate the effectiveness of the proposed approach.

In chapter 4, Continuous-time Markov Chain metabolic flux analysis method is adopted to central processes of the metabolism including the pentose phosphate pathway, glycolysis, and tricarboxylic acid (TCA) cycle in order to indicate that the approach can be used in a large and complex biological system network. The Monte Carlo method was used to solve the problem that not all the reaction rate constants k can be available.

In chapter 5, MFA is applied to biological models related to metabolic diseases. The insulin signaling pathway was selected as the model. The results showed that multiple homeostasis problems should be taken into consideration in the dynamic process, while multiple homeostasis in the insulin receptor is proposed to reveal the cause of insulin resistance in Type 2 diabetes through the reaction dynamic model.

In chapter 6, meta-analysis is further examined to gene expression during heart regeneration in Danio rerio. The results might play an important role in guiding future research on heart regeneration of zebrafish. The analysis revealed several genes that have not been linked to the heart regeneration process and might be useful for rebuilding cardiac function in zebrafish.

The overall research conclusion is summarized in chapter 7.

Chapter 2 Literature Review

2.1 Systems biology, metabolomics

In 1953, Watson and Frances [Wat53] found the double-helix structure of DNA and the life science research stepped into the Molecular Biology stage [Che00]. Molecular biology uses the reductionism method. The basic pattern of the reductionism method firstly is to divide the complicated things into multiple small basis on certain principles. Then, to further divide the small parts to smaller sub-component parts until the researchers can make a strict and thorough analysis of the small components. Finally, to understand the whole system based on the knowledge of the component parts. Molecular biology makes the discovery of a complicated life web possible. Once with some knowledge of this complicated process, the reductionism method has shown significant limitations facing complicated problems. The research on a single gene, individual metabolic pathways or an individual life phenomenon is difficult to provide researchers enough information to obtain entire knowledge of life [Con85]. At the same time, with the limitation of versions, the reductionism research is also likely to be puzzled by the disguise and to lead to a misunderstanding.

The development of biology needs to be driven by a new paradigm, and systems biology is a holistic approach for the latest biology research. Systems biology is a method within systems science. Systems biology is a qualitative research which regards biology as a whole system instead of many isolated parts. Systems biology studies the systematical behavior, correlation and dynamic characteristics of all

components of functional life systems by new technologies across many subjects. Also, it further demonstrates the basic rules of life system controls and designs. Systems biology enables us not only to understand all the components and the dynamic relations of complicated life systems, but also to predict future behavior when the system is stimulated and puzzled by the outside world [Ekd85].

The research on systems biology is mainly based on the following different aspects:

1. To realize the structure of a system, such as gene regulation and biochemical networks;

2. To study the behavior of a system, analyze the system dynamics and predict its behavior with the development of a theory and models;

3. To examine how to control the system and the mechanisms used by the system to control the cell state;

4. To understand how to design the system.

According to the theory, systems biology includes designing, improving and rebuilding the biology in order to control the system, as well as being the overall research target. It shows the following characteristics:

1. The research can start by the overall level. Systems biology gathers all the elements of biological systems, like gene, mRNA, protein, and protein interactions, in order to study the relation between these elements in response to a biological or genetic structure disturbance. Then the information in different levels can be integrated, and the behavior of biological systems in any conditions can be described. A new nature of biological systems by biological decoration or medicals

can be predicted in the future.

2. In order to put the emphasis on the use of information method, systems biology makes use of signal and systematical methods to research the dynamic process inside and between cells.

3. The modeling analysis methods adopted in practical application are usually mathematical models and simulations.

However, the basic premise of the development of systems biology is the large amount of omics data. Apart from the well-known genomics sciences, transcriptome and phenome in the omics data [Bin10], metabolome and fluxome have become increasingly important.

Metabonomics is corresponding to metabolome. It focuses on the Qualitative and quantitative to all the metabolites in the target systems. It focuses on the transformation of the whole metabolome. Compared with the information of genomics sciences, transcriptome and phenome, the data of fluxome have the self-advantages. Because the distribution of metabolic flux describes the physiological state of organisms directly. It is the result of comprehensive actions in genetic, metabolic regulation and environmental factors. To analyze the cell comprehensively with the combination of fluxome data, other omics information (like genomics sciences, transcriptome and phenome) has become a very important field in current and future research. MFA is the most effective method to make the quantitative analysis of microbial metabolic system in the steady state. It is suitable to determine the metabolic flux distribution in the network, analyze the

Metabolic capacity of the system and evaluate the effect of genes and environment to the cells accurately. It is a very important research fields on system biology on the basis of group index analysis, by means of high flux detection and data processing and aiming at information modeling and system integration. It reflects the changes of cells and tissue metabolism in the response to the outside stimulations and genetic modification. It is of great significance to depict the physiological characteristics of biological systems and assess the effect of heredity and environment to the cells. Not only the development of system biology difficult to do without the flow of omics, but also the rise and the development of metabolic engineering has raised more and more requirements of the flow of omics, which is regarded as its basic.

Metabolic engineering is an applied subject to systematically analyze the cell metabolism in cell metabolic pathways by molecular biology principle and design the cell metabolic pathway and genetic modification by recombinant DNA technology to achieve the cell trait transformation. In 1991, the project makes genetic manipulation and further modify the activity of cells by the enzyme reaction, substance transportation and regulatory of recombinant DNA technology to cells was described as the terms as metabolic engineering by Bailey [Bai91]. This is considered as a turning point of metabolic engineering to systematical subject. And then, the wide research of metabolic engineering has started. The new point of metabolic engineering is to focus on the combination of metabolic pathway instead of a single response. Therefore, it must study the intact

biochemical reaction network and put emphasis on the thermodynamic feasibility, metabolic flux and controlling of pathways and target products.

For the newly developed metabolic engineering, metabolic flux analysis is an experimental fluxomic techniques which is in a period of vigorous development and can provide very important supporting for future metabolomics and the system biology combined by the other omics.

The quantitative characterization of metabolic flux [Ben94] in every pathway of inner cells is the important goal of cell biology and metabolic engineering especially has guiding significance for substrate to effectively transfer to target products in the large-scale fermentation production of bioactive molecules.

Metabolic flux analysis is a method to analyze the metabolic pathway flow. On the basis of extracellular substance, the flow points of intracellular materials by the biochemical reactions can be calculated. Comparatively, the flow is obtained by the spectrophotometry in the uptake ratio and the secretion rate of substrate in the cells. The final task of metabolic flux analysis is to draw a complete metabolic process flow diagram which includes the complete metabolic pathway of all the relevant reaction and the flow distribution of metabolites in stable status.

Apart from the quantitative description of the various metabolic reactions, there are following roles:

1. The specific mutations used in the cell [Elo02] culture or changing the operation conditions directly may lead to the changes of the flow distributions in the nodes. Through comparing the changes, it helps to confirm that the node is

either hard stems or soft terms. Generally speaking, the flow distributions of node in the hard stems will not have big change, but the flow distributions of the node in the soft stems will adjust according to the changes of the environment and the cellular regulation. So, the accurate confirmation of nodes in the hard terms and soft terms in the metabolic network is very important to design of the pathway operations.

2. Choose the distinguish way: the chemical measurement of the biochemical reactions is the basis of the cell metabolic flux analysis, but the accurate chemical measurement needs the detailed background knowledge of the related biochemical reactions. However, the molecular mechanism of many biochemical reactions in the microbial cell is not clear. Metabolic flux analysis can help people to identify that which way is the best way to generate the optimal target product and which choices in material balance is impossible, then the accuracy and effectiveness of the target ways can be determined.

3. Non-determination extracellular substance flow calculation：Compared with the number of intracellular flows only obtained by calculation, sometimes the number of the measurable extracellular flow is very small. At this status, the flow section founded by the formal experimental results can directly calculate the measurement of material flow. For example, the generate rate of the target product relative to the by-products can be calculated by the chemical and specific metabolic flux.

To study the metabolic flux systematically, the control mechanism has three basic

processes:

1. Firstly, a measurement should be found to calculate the flow with the search methods as many as possible [Env09]. To achieve this, some simple material balance often can be made by the measurement of the concentration in the extracellular metabolites should be made. To be emphasized, the metabolic flux of the metabolic pathway doesn't equal to one or more enzyme activity in the pathway. In fact, the enzymatic analysis does not provide the real metabolic flux information in the pathways, unless the corresponding enzymes are existing and alive under the vitro analysis conditions. In the metabolic analysis, the enzymatic analysis often shows the similar order of magnitude about the power flow wrongly, and then leads to the incorrect conclusions.

2. Secondly, the ways that commonly used to apply a known perturbation in the biochemical metabolic network is to determine the pathway metabolic flux when the system transfers from a loose status to the new stable status. They are promoter induction; the substrate adding pulse and specific carbon source to eliminate or change the physical factors. Though any useful perturbation to metabolic flux is acceptable, the perturbation should be located in the enzyme molecule close to the way nodes. One perturbation can provide the information of many nodes. This is very important to the necessary minimum experiment amount of the accurate description on the control structure of metabolic network.

3. The results of metabolic fluxes perturbations should be analyzed systematically. If the perturbations of the metabolic flux difficult to make

observable effects to the downstream metabolic fluxes, the effect of this node to the upstream disturbance is regarded as rigid. Otherwise, it is flexible. In the rigid node, trying to influence the downstream metabolic flux by changing the activities of the upstream enzymes is meaningless. Fluxes are not observable effects, then you can think it node response to upstream disturbance is rigid, the opposite is called flexible, in node rigidity, trying to influence by changing the upstream enzymes downstream metabolic flux is futile.

2.2 Metabolic Flux Analysis

2.2.1 Quantitative Metabolic Flux Analysis

Quantitative metabolic flux analysis can mainly be divided into two parts: the MFA based on metrological balance and the MFA based on 13-Carbon Labeling. The former MFA is represented by flux balance analysis (FBA). FBA is a traditional method based on the mass conservation principle used to analyze the distribution of equilibrium metabolic flux. The data mainly comes from the measurement of extracellular flux. The ^{13}C MFA is a more powerful quantitative method. Carbon labeling quantitative measurement provides a large amount of input information and is the only exact analysis method used in metabolic flux distributions. Although both methods are different in principles, facts and results, the analysis of the microbial metabolic flux by MFA uniformly follows systems biology. When MFA is used to analyze the process of microbial metabolic flux analysis under the condition of steady state, it uniformly follows traditional systems biology methodology.

Firstly, it is found that the system models based on the qualitative biology knowledg. At the beginning, the model may not fully describe the biological system, as a serious of parameters in the model need to be further adjusted. Based on the system models, the biological system by computers is simulated (which are also called “dry” experiment). Due to the cause of the parameters in the models, the results of the early biological experiment have big difference with the real experiments (which are called “wet” experiment). Then the appropriateness between simulated results and the real results can be examined based on the constant adjustments of the parameters in the models. This process continues until the range error for the “dry” and “wet” experiments are acceptable. At this time, “dry” and “wet” experiments can be regarded as anastomosing. The adjusted models can reflect the nature of biological systems relatively true.

The target of system biology analysis methods is to gain the quantitative knowledge of system biology based on the qualitative knowledge. And then the biology systems can be more comprehensive and deep explain and discover the meaningful biological phenomena. Particularly in MFA, the qualitative biology knowledge mainly included the construction of metabolic network, constraint relations of metabolic fluxes and the choice of metabolic targets. Some of them can be obtained by books, some others, however, don’t have the information on hand and need to be improved by the analysis, like the metabolic target is one assumption of cell metabolic behaviors and whether the assumption is valid or not still need to be confirmed.

The mathematical models used in the analysis included the Metering balance equations founded by metabolic networks in Flux Balance Analysis (FBA) and isotopomer balance equations founded in 13C MFA. These models reflect the linear relations of fluxes in the network, defined the flux space of surface metabolic status set and the non-linear relations between the flux distribution and carbon isotope distribution in the network.

The parameters to be adjusted in the models are the specific values of flux distribution. The target of precise quantitative analysis in MFA is to give a flux analysis diagram which is fully approximate to the real situations. Flux analysis diagram is integrative determined by the gene regulation inside cells and external environment. It is one of the cell phenotypes and describes the unique physiological states of the cells corresponding to a particular solution in the flux space. In the basis of models, the 13C MFA simulated the Carbon Labeling Experiment in the computers. FBA optimized the object function in the flux space defined by metering balance and achieve the distribution of isotopomer and flux in the stable situation. This is regarded as dry test in the MFA, while the wet test is that the cell cultivation and carbon labeling are included.

Metering metabolic flux analysis developed from the late 80s in last century. At that time, the flux analysis is based on the metering balance model of the metabolic networks (which is called Stoichiomietric MFA). It took the extracellular metabolic flux as the main input resources. And then to conclude the distribution of intracellular metabolic flux. The flux distribution of the whole metabolic

network can be determined by the stoichiometry of the actions in the metabolic pathways, the substrates measured in the experiments, the rate of the products and the cellular component.

The method has been significant developed and is the basic method of MFA in the early stage. In 1993, Vallino [Val93] firstly described the full central metabolic as the metabolic flux map in the method. The system founded by this method is always under-determined. Equations always exist a certain freedom and difficult to certain all the intracellular flux uniquely. And often need to introduce the objective function to do some assumptions of biochemical systems and regard the assumption as the optimization objectives. The common assumptions are flux balance, minimum ATP production rate, maximum growth rate and so on. These measures represented by flux balance analysis are constraint-based method. It has lots advantages, like the relatively simple experiments, the strong operability, the flexible methods, relatively perfect theory, strong analysis capability of metabolic network and suitable to analyze the mass scale network. And now it is still the hot spot of research.

The introduction of metabolic target is not only the performance of the flexible method and the strong analysis ability, but also the makes the query of the rationality to implement the methods. Because the introduction of objective function is the specific assumption. Whether the assumption is true or not still need to be confirmed. At least, the energy balance has proved to be not true in some metabolic network. And the pure measurement method also difficult to deal with

the empty loop, bidirectional response, parallel reactions and exchange fluxes.

2.2.2 Measurement matrix analysis method

In the metabolic flux analysis methods, the metering Metrix analysis chart is one of the earliest methods based on metering models. It is based on the pseudo-steady state hypothesis, which is also regarded as mass balance approach. Based on the stoichiometry of the actions in the metabolic pathways, the substrates measured in the experiments, the rate of the products and the cellular component helps to determine the flux distribution in the whole metabolic network.

Assume *J* reactions happened in the cells with *N* substrates *s*, M metabolic products *p* and *L* intracellular variable quantity *x* (include intermediate metabolite and biological macromolecular substance), the measurement equation can be written as:

$$\mathrm{A} \cdot \mathrm{s} + \mathrm{B} \cdot \mathrm{p} + \mathrm{K} \cdot \mathrm{x} = 0 \tag{2.1}$$

Here, A, B, K represent the matrixes including all measurement parameters, assume the rate vector of *J* reactions is *r* (mols/g/h), substrate utilization rate and product formation rate are as follow:

$$r_s = -\mathrm{A} \cdot r^T$$

$$r_P = -\mathrm{B} \cdot r^T \tag{2.2}$$

The formation rate based on the intracellular variable is:

$$\mathrm{R} = \mathrm{K} \cdot r^T - U \cdot X \tag{2.3}$$

Where X (mol/g) is intracellular component concentration, when the cell growth, $-U \cdot X$ is the dilution of intracellular component as the matrix expands, U

represents the formation rate.

In metabolic flux analysis, with the components of rate vector *r*, combine r_s, r_P and R to obtain the intracellular flux distribution under certain conditions. According to the quasi steady state assumption, *C* metabolic intermediates bring *C* constraint conditions on rate. So

$$K_C \cdot r^T = 0 \tag{2.4}$$

Where *r* is the submatrix when stoichiometric coefficients is 0, then degree of freedom is

$$\mathrm{F} = \mathrm{J} - \mathrm{C} \tag{2.5}$$

Some consume and generating rate of extracellular can be measured. By setting the speed rate as *m*, if $\mathrm{m} > -\mathrm{F}$, the whole flux distribution can be confirmed. For the system under the condition $\mathrm{m} < \mathrm{F}$, due to the number of the constrain equations about the model is less than the unknown fluxes, the solution to the model fluxes is uncertain. And the system is regarded as the indefinite system. For the system with the condition that $\mathrm{m} = -\mathrm{F}$, which exists the only solution, it is called chengting system. For the system $\mathrm{m} > \mathrm{F}$, the number of equations is more than the number of unknown fluxes, and it not only has the only solutions, but also can be verified, this system is regarded as over-determined system. With the different systems, there are different solution methods to be chosen to make the calculation of the metabolic fluxes. To calculate the uncertain systems, an optimization goal should be set. With seeking the target optimization, the valuations of the modeling fluxes distribution can be obtained. The matrix

transformation method can be directly used to make the solution of changing system. And for the solution f over-determined system, the least square method can be usually taken.

There are many important issues happened in the metering method development process. In 1984, Papoutsakis started to calculate the maximum theoretical outputs by the linear programming method [Pap84]. Then Fell and Small started to research the generation of fat by linear programming method in 1986 [Fel86].

After stepping into the 90s, the FBA started to develop quickly. Savinell and Palsson made full analysis for FBA, and then the theoretical basis had been founded [Sav92]. Varma and Palsson used FBA to describe the nature of colibacillus in 1993 [Var93]. Pramanik and Palsson studied the relationship of the generating rate of cells and the concentration of Biomass by FBA [Pra97]. In 2000, Edwards and Palsson made gene knockout, phaseplane and robustness analysis to colibacillus; and Schilling et al combined FBA and extreme pathway analysis [Edw00]. In 2001, Burgard and Maranas studied the performance limitations of colibacillus and got the minimum response set. And Covert et al added the control restriction to FBA models [Bur01]. Mahadevan developed the dynamic flux balance analysis to study the dynamic changes in regulation of gene expression and introduced the non-linear energy balance restriction to the FBA to quantify the concentration of the intracellular metabolite, but it is difficult to guarantee that the solution is the global optimum in 2002 [Mah02].

The development direction of next generation of FBA including:

1. To research the new target equation to study the ability, limitation and robustness of the microbial metabolism fully and deeply;

2. To introduce more biological information as the knowledge restriction, such as control restriction and thermal dynamic constraint;

3. To develop the dynamic FBA and analyze the cell metabolic behaviors in the unstable status.

2.3 ^{13}C Metabolic Flux Analysis

2.3.1 Background of ^{13}C MFA

Among the MFA method, the ^{13}C MFA based on carbon isotope labeling experiment is more accurate. Compared with the calculation method, the main difference is the ^{13}C MFA based on the carbon labeling experiment.

Carbon isotope labeling experiment regards the specific substrate with ^{13}C labeled, like the compound of [1-^{13}C] glucose and natural glucosamine, as the uptake of the biological system. After a period of internal response and the whole system become stable, then the marked carbon atoms will be distributed to all positions in the metabolic network. Then some of the isotope labeling can be calculated by the magnetic flotation resonate and mass spectrum. These data can bring restraint information for the quantitative analysis of the intracellular flux in the metabolic network. The flux in the biological can be analyzed accurately based on the information. Until now, ^{13}C MFA is the only common and accurate flux analysis method does not rely on the metabolic network and has been widely applied and researched among the world. Compared with quantitative method, this is a more

powerful and reliable flux quantitative analysis method, as well as a more complicated method.

Firstly, the experiment operation is very complicated. Microbial cells need to be crushed. The giant molecule needs to be hydrolyzed. And a large number of measured specimens need to be prepared. The carbon labeling experiment takes long time to continuous culture.

Secondly, the specimens used in the carbon labeling experiment are very expensive, for example, one gram of [1-^{13}C] glucose takes about 100 dollars in the current market. And there are also many preconditions of the experiences. The cells should achieve the stable labeling status of the metabolic and isotope. And the isotope labelling and experimental measurement should be supposed as no influence on the cellular metabolism.

And the calculation is more complicated; the internal mathematical model often involves thousands of dimension matrix operations and often needs to iterate and random explore. The calculation usually takes a few hours or even few days. ^{13}C MFA, however, develops extremely quickly in the recent years. The further development of experimental techniques and calculation methods relieve the complexity of the operations, for example, using the more mature modern mature sample preparation procedures and high precision measuring instrument can shorten the experimental time from few weeks to several days. And the proliferation of the powerful data analysis system and hardware computing ability reduced the difficulty of data analysis. All in all, the development of ^{13}C MFA has

experienced the following several major milestones.

In 1982, Blun and Stein applied the ^{14}C labeled metabolic flux analysis to Tetrahymena cells. At that time, the analysis method is extremely time-consuming [Blu82]. Malloy made the newly established work in the isotope analysis when they study the citric acid cycle activity inside the heart [Mal88]. Zupke and Stephanopoulos finished the first metabolic flux analysis by the common mathematical modeling method in 1995 [Zup95]. Marx generated the first labeling data set including 25 pieces of 1HNMR measurement data [Mar96]. Wiechert [Wie97] fully described the common statistical analysis proposal with the ^{1}H NMR measured data for ^{13}C metabolic flux analysis. Schmidt [Sch99] used parameter fitness method to evaluate the data of $^{1}H/^{13}C$ MFA 2D COSY NMR in colibacillus. In 1999, Sauer and Bailey [Sau99] make study on the most promising direction--cell energy metabolism in ^{13}C metabolic flux analysis. The introduction of mass spectrum data in ^{13}C metabolic flux analysis brought the new mathematical problems. Park studied covering flux among the C.glutamicum by a widely-used but unusual modeling method [Par00].

The common application of ^{13}C MFA is to compare the variable physiological status by ^{13}C MFA. And Graaf [Gra99] did the responding work in Zymomonas mobilis. Christensen and Nielsen [Chr00] put forward to a universal and easily calculated ^{13}C MFA frame for Peniciliumchrysogenum. Petersen [Pet00] firstly combined the universal modeling methods and experimental design methods to confirm four covering fluxes in C.glutamicum. Winden [Win02] put forward the

bondomer labeling experiment-based ^{13}C MFA, which equals to isotopomer modeling to save the calculation time.

John [Joh03] applied the microtiter plate technology to lysine Corynebacterium ammoniagenes biological response system and realized the ^{13}C MFA to process in the micro scale cultivation conditions. Drysch [Dry03] firstly combined the biosensors reactor technology with ^{13}C MFA with the utilization of Corynebacterium ammoniagenes in the medium scale production of L-lysine. Fischer [Fis04] analyzed the influence of the substrates with different labeling models to the calculation of colibacillus central carbon metabolism flux by microtiter plate technology and shake flask culture parallel test in 2004.

MFA uses the main chemical reactions' stoichiometric model and material balance (black box model) of intracellular to calculate the intracellular metabolic flux, so it is called metrology MFA. Metrology MFA's metrological parameter mainly includes the substrate uptake rate, measurement parameters, product formation rate, biomass and the release of CO_2, etc.

Yet the metrology MFA shows many limitations when deal with the practical problems: Firstly, the limitation features on the energy balance analysis. Energy balance analysis helps to identify the cells cofactors (such as NADH and NADPH) and the reaction of ATP's production and consumption. To fully determine the energy-related reaction is very difficult as the calculated flux is not accurate. Secondly, there are many invalid loops in prokaryotes. It leads these leads to many difficulties for metrology MFA to deal with energy balance. What's more,

metrology MFA is difficult to determine reversible reaction, two-way, covering and the net flux parallel reactions in cells. Then ^{13}C markers (^{13}C labeling experiments, CLEs) –based ^{13}C metabolic flux analysis (^{13}C metabolic flux analysis, ^{13}C MFA) is developed based on the above limitations.

2.3.2 Principles and Methods of ^{13}C MFA

Based on the econometric model and material balance of Metrology MFA, ^{13}C MFA can establish a balance equation for a single carbon atom by tagging the substrate and analyzing the intracellular metabolites tag mode by using NMR and MS, which exclude cofactor balance equation.

In addition, because many single carbon atom balance equations are available, you can get an over-determined system. The equation's redundancy can estimate the unknown flux more accurately.

So, ^{13}C MFA is therefore a more accurate calculation method of unknown flux, and now has become an important tool for analyzing complex metabolic network flux. The intracellular flux can be solved by measuring the extracellular and intracellular flux tag information.

To conduct ^{13}C MFA includes the following three important constituents, they are CLEs, data measurement and the evaluation of CLEs. The following will be a brief introduction of the three parts.

1. ^{13}C labeling experiment

In the ^{13}C labeling experiments, ^{13}C-labeled substrate (usually use glucose) is introduced into the bioreactor system. Along with the biochemical reaction's

generation and after chemical rearrangement, bond breaking and new bond's generation, ^{13}C atoms will distribute in the metabolic networks, intracellular metabolite and biomass composition.

With the consumption of glucose, labeled carbon's isotopic abundance in Metabolite Library increases, and until it can be detected by NMR or MS, the samples are taken out to further analyze. Here are some tips to do the ^{13}C labeling experiments:

- Steady-state assumption is an important prerequisite for the MFA. In the 13C labeling experiments, not only to ensure the system in metabolic homeostasis, but also the system in isotopic equilibrium; the mark balance of biomass components (proteins, nucleic acids, lipids) are taken into consideration, because the intermediary metabolism was usually not directly measured, what the chemical compounds synthesized by precursor is measured.

- It must be very cautious to select the labeled substrate. In the glucose-substrate reaction system, the uniformly labeled substrates are [1-13C] glucose and uniformed tagged glucose. The labeled substrate's price is usually around $100 per gram, but the other specific marked glucose's prices are even more expensive. So the labeled substrate's combinations should be designed reasonably, to receive more flux in a relatively lower experiment cost.

- In order to reduce the substrate's expenses in the experiments, the reactor's volume should be small enough. But if the reactor's volume is too small, the system's stability will be very hard to reach. So currently, the reactor's volume

is usually 300 to 1000ml.

- In terms of sample processing, proper sample processing methods of different metabolites should be chosen to ensure that the compound structure would not be damaged.

2. Measurement technology platform

Nuclear magnetic resonance (NMR) technology is an analysis technology to identify the compound structure of radiofrequency radiation absorption spectra by nucleus in high magnetic field. Because the NMR technology can be used in the structure determination of amorphous material which X-ray crystallography difficult to determine, it plays very important role in life science, medicine and materials science research. There are three methods commonly used in life science fields: hydrogen spectrum (^{1}H NMR), carbon (^{13}C NMR) and phosphorus spectrum (^{31}P NMR) and are used in the fluid extract (or organization) and in-vivo analysis.

For example, ^{1}H NMR the biological specimens prepared directly on the sample testing. The ^{1}H NMR spectrum peak correspond with hydrogen atom of each compound in the samples and compared with the standard NMR or based on certain rules, then the chemical composition of the metabolites can be identified. And the relative strength of the signals can reflect the relative content of each component.

Mass-spectrometric technique is to sort the ionized atoms, molecules or molecular

fragments as qualified ratio (mPe) into atlas and based on the atlas to make the qualitative or quantitative analysis for all kinds of materials. If the pretreatment biological samples were added to a mass spectrometer, the corresponding metabolite atlas would be obtained after gasification, ionization, accelerated separation and detection. Each peak corresponds to molecular weight in the atlas, after the further detection and analysis, chemical composition and semi-quantitative relationship can be diagnosed partly.

As ^{13}C MFA measurement technology platform, nuclear magnetic resonance (NMR) and mass spectrometry techniques have their respective advantages and disadvantages, and they keep overcoming the shortcomings to develop. Details see the literature.

3. Evaluation to CLEs

The establishment of model and simulation is the core of CLEs data evaluation. In the simulation process, assuming that the intracellular flux is known, the steady state distribution of the isotope labeling the isomers can be calculated according to the assumed intercompared and substrate composition. Based on CLEs simulation， the flux set can be revised systematically by the corresponding assessment algorithm. The widely used algorithm routine and the detailed process as followed:

- Set some flux distribution on the metabolic network and establish the stoichiometric balance equation;

- Mark CLEs simulation for substrate isotope labeling isomer based on the above flux distribution;
- Insert the given flux into the system and calculate the measured value according to substrate isotope labeling isomer distribution results;
- Calculate the errors between the measured value and simulative predicted value;

Revise the given flux systematically by optimization algorithm based on the results obtain. Since between the biological system and the measured values, there always exists severe “noise”, and the transfer of the errors between them will lead to the results which seem accurate but meaningless. It is very important to make statistical analysis to the estimated flux after parameter fitting procedure.

2.4 Development of MFA

2.4.1 Experiment

1. Impact on flux analysis by substrate labeling pattern

^{13}C labeling experiments have very strong work strength and high cost. So how to input lowest human and material resources to maximize unknown flux information at most caused the great interest of the experimental people. When choosing the approaches, measurement method, and multiple substrate labeling model to make the optimal design for the experiments sometimes can attain the result with half effort. The labeling patens of substrate can divide to the following types:

- Unlabeled substrate is always used as a component of mixed substrate in a tag experiment, since its price is relatively cheap;
- Specifically, label substrate is to tag a point specifically in the substrate carbon atoms, like [1-^{13}C]-glucose refers to a glucose which labeled ^{13}C in one carbon atom;
- Multipling specifically labeled substrate is to label multiple points in the substrate carbon atoms, like [5, 6-$^{13}C_2$] - glucose refers to a glucose labeled ^{13}C in the fifth and sixth carbon atoms at the same time. The glucose with labeled pattern is more expensive;
- The uniformly labeled substrates, like [U-^{13}C] - glucose refers to a glucose labeled every carbon atom.

Obviously, making the experiment with uniformly labeled substrates will not gain more flux information, just as with the unlabeled substrate pattern. So the above two labeled substrate pattern always mixedly use with other patterns in the ^{13}C experiments. Flux analysis experiment with different labeled substrate patterns are firstly suggested by Szyperski. This pattern comes from C-C bond labeling experiment (BLE).

Some researcher like Bender [Ben13] used covariance matrix to optimize the labeled substrate patterns, and make a conclusion that the best labeled substrate patterns to gain the most reliable information of the Synechocystis sp. PCC6803 center carbon metabolism flux is to mix 70 % unlabeled substrate, 10% [U-^{13}C]-

glucose and 20% [1-^{13}C]- glucose.

Fisher et.al used many different labeled substrate patterns to analyze the E. coli logarithmic phase center carbon flux distributions and analyze the relationship between the labeled substrate patterns and flux determination ways by experimental method at the first time. The results show: mixing [U-^{13}C]- glucose with the unlabeled glucose is suitable to parse the PEP pathway flux and C - C key fracture ways of metabolic flux caused by the flux exchange; and as the [1-^{13}C]-glucose as the substrate will bring a higher accuracy in the calculation of the central carbon metabolism upstream path, especially oxidative phosphorylation and ED pathway flux; the mixture of [U-^{13}C]-glucose and [1-^{13}C]-glucose is suitable to determine the metabolic network flux.

2. Microtiter well-plate technology

There are two methods to achieve high throughput screening: one is miniaturization, the other is automation. Based on above two tools, the high-throughput experiment is developed as microtiter plate technology. Microtiter well-plate has borne hole 896 and 384, 1536 plate hole titration and the supporting facilities are underdeveloped. The volume of microtiter well-plate technology is always under ml level. Each sample hole usually adopts the way of silicone membrane to ventilate individually by electrolysis. Under this mode, the oxygen transfer rate can reach K_La as 0. 042s^{-1}. Injecting stainless steel tactile to sample hole approves the medium and low mixture efficiency of the microtiter plate. At

the same time, injecting all kinds of sensors in this technology achieve many parameters (like pH value and dissolved oxygen) to achieve online monitoring. Meza [Mez12] successfully applied microtiter well-plate technology to the biological reaction system which produces lysine and Corynebacterium ammoniagenes and achieved the ^{13}C MFA to proceed in micro scale cultivation conditions. Fischer [Fis05] and the others made a parallel experiment by the microtiter well-plate technology and shake flask culture to analyze the influence of the different label mode of *e.coli* center carbon substrates metabolic flux calculation. The results under these two conditions are very similar. Microtiter well-plate technology which mixed miniaturization and automation was introduced to the metabolic engineering field. Then ^{13}C MFA can be proceeded by miniaturization and high flux, and this action also reduced the in-substrate consumption expenses of ^{13}C MFA and won the experiment time. Meanwhile, the technology can provide multiple parallel experiments. The microtiter well-plate technology shows great superiority in high-throughput screening experiments, like mass screening of the mutant strains which metabolic engineering, drug discovery and systems biology concerned most at the present. And the research of corresponding supporting facilities, such as porous suction sampler, centrifuge and concentrator, will provide microtiter well-plate technology vast application space.

3. Sensor reactor technology

The application of ^{13}C experiment in laboratory scale has become increasingly mature. The technology, however, impeded itself to be a standard technique when

it is applied to the biological reaction system in mass production conditions because of its own disadvantages. Firstly, the premise of ^{13}C MFA is the markers balance, which limits the application applied in the stable the steady or pseudosteady conditions [Fri09]. Secondly, once the biomass is marked, the complex results and measurement noise makes the further mark unable to process. In addition, the tagged substrates' cost made the small-scale labeling experiment more attractive than industrial application. The reasons above largely limited the application of ^{13}C MFA in broader application. Until the biological sensor reactor technology was introduced into biological mass cultivation system, ^{13}C MFA was applied to the industrial level. The sensor reactor (volume as 1L) connects to the mass-production-scale production reactor (volume as 160L) in slaved mode. When the labeling experiments are required, few culture mediums in the production reactor will be imported into the sensor reactors and keep two reactors in paralleled train status by master-slave mode control at any time. The labeling experiments happen in the sensor reactors. And when the labeling experiments are finished and the thallus will be collected for ^{13}C MFA, the sensor reactor will return to the initial state and get ready for the next labeling experiment. When Kadirkamanathan [Kad06] and other scientists used the Corynebacterium ammoniagenes to make the L-lysine Chinese style scale production, they combined the biological sensor reaction and ^{13}C MFA at the first time, investigated the metabolic flux distribution center carbon of the Chinese style scale fermenter lysine production bacteria C. glutamicum MH20-2213 in 3-period exponential phase and analyzed the change

of anaplerotic node metabolism flow distribution with the increase of the lysine biosynthesis rate. By analyzing the consistency of the fermentation behavior between sensor reactors and production reactors, it is approved that ^{13}C MFA in biological sensing reactor can well reflect the metabolic flow reaction system. And the ^{13}C MFA can be used to industrial scale. Since currently the ^{13}C MFA applied sensor reactor technology to large-scale biological reaction system limits in thallus exponential phase. The real industrial fermentation involves the growth of bacteria at different stages. So it is still very challengeable to mix the biological reaction and ^{13}C MFA to apply into every stage of wide-range liquid fermentation.

2.4.2 Data analysis

Isotopomer is the combination of the word “isotope” and “isomer” and refers to a labeling status a metabolite may face. For example, when considering the labeling status of every carbon atom, 2n different isotope labeling isomers will be produces in a metabolite the n carbon atoms. The distribution of isotope labeling isomer is reflected by the percentage of the labeling isomer in the metabolites’ library [Fuh05]. Obviously, the sum of the percentages of labeling isomer in specific isotope labeling isomer library is 1.

The concept of cumomer is developed based on the isotope labeling isomers and refers to the collection of isotope labeling isomer with specific markers. Cumomer refers to the sum of isotope labeling isomer with specific markers. For example, isotope labeling isomer group M # 11 X refers to all the tagged isomers of which the first and second carbon atoms in the molecule M are marked and belongs to a

specific labeling carbon atom marked as 1, otherwise as 0. The corresponding Cumomer scores m_{11x} can be calculated by m_{110} plus m_{111}. Under this definition, a given isotope labeling isomer can be introduced into different cumomer labeling models and makes contribution to the cumomer scores for different labeling models.

Functional genomics study is to analyze the gene function comprehensively based on genomics or system level by using the information provided by genome structure and the application of new experimental methods. Below are the steps to determine the cell phenotype from the genome information: identification of genes, enzyme catalytic reaction column into metabolic map and kinetics and regulatory analysis to actual phenotype [Fre12]. Currently, the genetic sequencing develops toward velocity and economization, and the judgment of gene functions has become easier by advanced comparative sequence. It is impossible to identify the functions of a large number of isolated genes existed in the existing in the genome by comparative sequence analysis. To identify the function of the isolated gene by elementary flux mode is a very effective method. Meanwhile, the determination of elementary flux mode is also essential to econometric model structural analysis. The functional genomics research needs to integrate every group no matter in genome, transcriptome, proteome, metabolites, flux or the related technologies to analyze the genome function in multi-level and comprehensively. Functional genomics research usually adopts system analysis methods, including providing rich mRNA transcription omics information analysis method, the proteome

information proteomics analysis method and the metabolic profiling analysis method which can reflect the despondence of steady-state metabolite levels to genetic, physiological and environmental disturbance. And the flux group information used as an important aspect of cell metabolic state also plays a decisive role in functional genomics research. Additionally, the phenomenon, the metabolic flux information changeable didn't cause any change in the corresponding metabolite levels and even in final production, usually happened in metabolic pathways. There must be some regulation mechanism working under this circumstance. The flux information as the sensitive parameters provides very important information to reveal the gene function. Therefore, the determination of metabolic flux become a very important part to the functional genomics research.

2.5 Metabolic homeostasis and insulin resistance

2.5.1 Mitochondria homeostasis

Mitochondria are the intracellular energy metabolism plants. Their function gradually declines with aging, suggesting that there may be a relationship between mitochondria and aging. The decline in mitochondrial activity and functional disorders can disrupt the maintenance of stem cells and promote accumulation of aging factors such as cellular aging and chronic inflammatory response [Sun16]. Mitochondrial DNA mutations are significantly increased in mitochondrial DNA polymerase mutant mice, and their aging rate is also significantly higher [Tri04]. Conversely, increasing the biosynthesis of new mitochondria can delay the

senescence of smooth muscle in mice and prevent the dystrophy that is exacerbated with age [Wen09]. In summary, the decline in mitochondria is likely to be one of the causes of aging in organisms. Although the normal function of mitochondria seems to be beneficial for aging, there is increasing evidence that mitochondrial damage or weak function can extend lifespan ([Liu05], [Dil02]). An in-depth study of this seemingly contradictory view suggests that weak damage to mitochondria causes a mitochondrial unfolded protein response (mtUPR), which in turn mediates the longevity of the organism [Dur11].

MtUPR is stress response of mitochondria. After mitochondrial damage, mtUPR induces the expression of mitochondrial chaperones and restores mitochondrial function to some extent; mtUPR also improves the metabolism and regulates the internal immune response to help cells survive as much as possible under conditions of imperfect mitochondrial function [Sch15]. mtUPR may have contributed to the longevity of the organism through these mechanisms.

2.5.2 Insulin homeostasis

Insulin is an important hormone regulating sugar metabolism in animals. Abnormalities in the insulin signaling pathway are the cause of type 2 diabetes. This pathway is also the first signal pathway that has been found to affect animal aging. From the 1980s to the 1990s, scientists have found that the lifespan of Caenorhabditis elegans, which lacks the insulin/insulin-like growth factor receptor (daf-2) or an important kinase (age-1) in the pathway, is 2 to 3 times longer than that of wild-type nematodes, and these longevous nematodes still maintain a young

phenotype for a long time after the wild type nematode shows an aging phenotype ([Ken93], [Kla83]). In these mutants, the insulin/insulin-like growth factor-1 signalling pathway is inhibited and the activity of the downstream transcription factor daf-16 is enhanced ([Lin97], [Kim97]). This activates the expression of many downstream genes and prolongs the nematodes' lifespan. Longevity of the insulin/insulin-like growth factor-1 signaling pathway is evolutionarily conserved, and its role is not limited to nematodes: appropriate inhibition of some components of this pathway in fruit flies, mice, and dogs can also present longevity effects [Ken05]. Even in humans, long-lived groups are sometimes accompanied by mutations in some pathway components, indicating that the life-regulating effects of this pathway are equally applicable to humans ([Koj04], [Paw09]).

2.5.3 Target of rapamycin and energy homeostasis

TOR is the main amino acid and nutrient receptor in the body, promoting anabolism and driving growth when food is enough. AMPK is another nutrient and energy receptor in the body that promotes catabolism and inhibits anabolism by sensing the increase in the ratio of AMP to ATP. Enhancing AMPK activity or inhibiting TOR activity in a variety of model animals can both extend their lifespan ([Apf04], [Ani08], [Kae05], [Jia04], [Har09]). TOR and AMPK functions antagonize and inhibit each other in vivo ([Hin15], [Bol02]). When TOR is inhibited and AMPK is activated, the synthesis of proteins, lipids and sugars in the whole cell is reduced, and the proteins, lipids and sugars stored and discarded in the cells are decomposed and reused. Respiration is enhanced, and the effect of

prolonging life is finally achieved. This pair of central metabolic regulatory molecules regulates aging of organisms and plays an important role in the aging process of organisms.

2.5.4 Cardiovascular insulin resistance

Insulin is a classic metabolic regulating hormone and plays an irreplaceable role in the regulation of glucose and in the lipid metabolism in the body. In addition to regulating the metabolism, insulin can also directly activate the cell "survival signal" to play a cardiovascular protective role, suggesting its importance in maintaining normal cardiovascular function. Additionally, clinical studies have shown that insulin resistance is the common pathological basis for major chronic diseases such as diabetes, coronary heart disease, hypertension and heart failure [XuY13]. Approximately 70% to 80% of deaths in patients with type 2 diabetes mellitus (T2DM) are caused by cardiovascular complications, suggesting an important role for insulin resistance and metabolic abnormalities in the occurrence and outcome of cardiovascular disease ([Che05], [Tan17], [Hab16]).

In the past, it was considered that insulin resistance, which presents in skeletal muscle, fat and liver tissue. It was characteristic of obesity and T2DM, and usually leads to systemic changes such as abnormal glucose tolerance, hyperinsulinemia, and hyperlipidemia [Mun07]. Recent studies have shown that patients with T2DM usually have cardiovascular insulin resistance ([Far16], [Bru14], [HaB16]). The main hallmarks of myocardial insulin resistance are decreased glucose uptake by cardiomyocytes stimulated by insulin, decreased phosphorylation of protein kinase

B (Akt), and decreased myocardial positive inotropic effects induced by insulin. Insulin resistance is mainly manifested by insulin stimulated Akt and endothelial nitric oxide synthase (eNOS) phosphorylation, decreased nitric oxide (NO) production in endothelial cells, and decreased vasodilation induced by insulin ([HaB16], [YuQ10]). Intravascular local insulin sensitivity reduction is not a direct cause of systemic metabolism and blood glucose disorders but is directly involved in the occurrence and development of diabetes-related cardiovascular disease.

2.6 Heart regeneration

The zebrafish (*Danio rerio*) has become a widely used vertebrate regeneration model due to its many regenerative organs such as the central nervous system, heart, kidney, and fins. It has been a well-developed experimental tool in recent years [Gem13]. Fin and heart regeneration are the two hottest areas of study, and there has been no small progress in the cellular and molecular mechanisms of regeneration of these two organs [Pos03]. Among them, the ability of the heart to regenerate is particularly important, and it has great guiding significance for the treatment of cardiovascular diseases in humans.

2.6.1 Zebrafish heart regeneration

As demonstrated by previous studies, zebrafish is one of the most important developmental and regenerative biological models due to its advantages of in vitro fertilization, fertility, embryonic transparency, in vitro development, and sequenced genomes [Gem13]. Over the past 25 years, hundreds of mutant fish

lines have been identified, and a variety of genetic tools that were first used in fruit flies and mice have been successfully applied to zebrafish. Compared with the mammalian heart, the zebrafish has a simpler heart structure with only one atrium and one ventricle, and it is smaller in size, approximately 1 cm^3, but its histological composition is very similar to that of other vertebrates. The study of the fish heart allows us to understand the developmental features of the more complex heart of other vertebrates from a simple perspective. The zebrafish model is fully utilized in cardiovascular development studies because zebrafish mutant embryos with defective blood circulation systems can still survive for 5 days after fertilization. Through genetic screening strategies, many essential genes that play an important role in cardiovascular development have been identified [Liu12].

Many organs of adult zebrafish have amazing regenerative capabilities, including all fins [Pos03], retina [Vih00], spinal cord [Bec97], telencephalon [Kro11] and kidney [Die11]. However, the mechanism regulating the regeneration process seems to be different in different organs. For example, the regeneration of fins is dependent on the formation of bud bases, a highly proliferating structure composed of dedifferentiated cells that can form all structures of the regenerated fins [Pfe15]. Conversely, regeneration of the telencephalon does not require the formation of a bud-like structure but requires activation of the progenitor cells in the telencephalon, which can be labeled with one of the target genes of the Notch signaling pathway, *her4.1* [Kro11]. In 2002, Poss and Keating discovered the strongest cardiac regenerative response in vertebrates to date [Pos02]. They found

that after removal of 20% of the ventricles, the zebrafish heart can still be effectively regenerated. After the resection, the zebrafish heart quickly forms a fibrin clot to stop the bleeding process. In the next few weeks, the fibrin clot is gradually replaced by new cardiomyocytes, eventually 30 days after the injury. The ventricle reached complete regeneration in 60 days [Pos02, Ray03]. Interestingly, cardiac regeneration does not occur in all bony fish, although in other species such as Carassius auratus and Devario aequipinnatus, the heart can be successfully regenerated [Gri14, Laf12], but only scar tissue can be formed in Oryzias latipies [Ito14]. Exploring the different responses of zebrafish and barley to heart damage is an effective way to identify essential factors in cardiac regeneration.

2.6.2 Zebrafish heart injury model

In the first decade of this field, the only damage model for zebrafish heart regeneration research was the ventricular resection. This method of injury involves the removal of a portion of the apex of the ventricle, and the degree of regeneration is measured by the extent to which the resected tissue is fully elongated [Pos02, Ray03]. One of the most important aspects of assessing regenerative capacity after excising the ventricular end of the zebrafish heart is the absence of scar tissue formation. Initially, there was a widely accepted view that myocardial scarring and regeneration were two mutually exclusive events [Sch04]. Since the excision of the ventricle is based on tissue removal rather than cell death, the heart itself does not need to clean up the debris of the dead cells. Thus, the zebrafish heart will not

form scar tissue after being removed. Whether the zebrafish heart can regenerate after more severe injuries has been an unknown event for many years.

As research progresses, a series of alternative injury models that induce tissue death have been established, and these models can better study the zebrafish heart regeneration process. Due to the small zebrafish heart size, it is not possible to use a more common method of injury such as coronary artery ligation, to induce myocardial infarction similarly to large animal hearts. The first damage model other than ventricular resection is the low temperature injury method [Cha11b, Gon11, Sch11]. This method first pre-cools the wire in liquid nitrogen and then the surface of the ventricle is touched with a liquid nitrogen pre-cooled wire to freeze a portion of the ventricle. Rapid freezing and thawing of the cells lead to tissue necrosis, and then the cells surrounding the necrotic area begin to enter the apoptosis program. As in the process of myocardial infarction in humans [Sar97], hypothermia causes rapid cardiomyocyte enucleation, and myofibrils are relatively unaffected within a few days [Gon11]. Compared to ventricular resection, injury by hypothermia induces a more severe apoptotic response, affecting all types of cells, including the epicardium, endocardium, and coronary arteries. In a short period of time after the injury, the immune cells infiltrate into the damaged area, fibrous tissue deposition occurs and myofibroblasts begin appearing wound site. Unlike in mammals, myocardial fibrosis in zebrafish is temporary, and scars are completely removed and replaced by cardiomyocytes within 3-4 months. In goldfish and squid, the same reaction occurs after burning the heart [Gri14, Laf12].

These results indicate that fibrosis itself does not inhibit the regeneration process. Compared with other injury models, the ventricular regeneration process after hypothermic injury was significantly slower, even though the size of the tissue removed during the injury process was similar. This delay may be due to the need to remove necrotic material in order to begin the regeneration process. Initial cryo-injury studies used copper wire to induce ventricular damage, and fibrous tissue scars were removed within 130 days for complete regeneration [Gon11]. Additional studies have used thicker copper wire to induce more severe damage, which was regenerated 180 days after injury [Hei15], suggesting that the extent of injury may affect the recovery time window. In addition, the use of platinum instead of copper in inducing low temperature damage also affects the size of the damaged area [Gon12]. Reconstituted hearts after hypothermic damage to the ventricles have different characteristics, such as enlarged ventricles, thickened damaged walls, and more rounded ventricular shapes. A recent study reported that in the heart of hypothermia, the primary layer showed incomplete regenerative capacity [Pfe17]. A phenomenon is not found in the regenerated heart after ventricular resection [Gup12]. The pumping efficiency of the entire heart is restored, but some areas of the ventricle are permanently affected after low temperature injury.

A third strategy for inducing cardiac damage relies on an inducible genetic system that specifically destroys cardiomyocytes through the expression of catalytic enzymes of toxins or cytotoxin metabolites. Transgenic fish lines that specifically

express bacterial nitroreductase (NTR) in ventricular myocytes are established. Nitroreductase alone is not toxic, but it can catalyze the conversion of the prodrug, metronidazole (MTZ) to its metabolite and induce cell death. Since MTZ can be added or removed from fish water at any time, this system has the flexibility to control the time and duration of injury [Cur07]. In subsequent studies, this technique was used to specifically disrupt ventricular myocytes in the developing zebrafish heart, revealing the ability of central atrial cells to acquire a ventricular phenotype during embryonic cardiac regeneration [Zha13b].

In order to study adult cardiac regeneration, a second system of genetically induced damage, the conditional expression of Diphtheria toxin A (DTA) in cardiomyocytes, was also established [Wan11]. Once DTA expression is induced, it promotes diffuse lesions that can cause damage to approximately 60% of cardiomyocytes. Although this large amount of cardiomyocyte damage can be tolerated by zebrafish, these animals begin to exhibit symptoms of heart failure, such as lethargy, convulsions, and low tolerance to exercise. Complete regeneration after myocardial injury can be achieved in about 30 days without scars, probably because the expression of DTA does not affect the endocardium and epicardium. Since myocardial damage specifically induces cardiomyocyte death, this approach is more like a precise, advanced cardiomyopathy model than a myocardial infarction, but this method is still a very good system for identifying factors related to cardiomyocyte proliferation. The genetic damage system can be used to analyze the cardiac regeneration process in zebrafish larvae and juvenile

individuals([Cur07], [Zha13b]), a perfect system for screening reproductive phenotypes using genetic and chemical methods [Cho13]. In order to better simulate the human myocardial infarction process, a hypoxia-reoxygenation model was also established [Par13]. Since the entire animal body is exposed to a hypoxic environment, this model induces not only damage to the heart, but other organs can also detect subsequent immune responses caused by the injury. Although this treatment induces apoptosis and proliferation in the heart, there is no damage visualized by histological analysis. After the hypoxia-reoxygenation process, the individual animal will have a transient cardiac function weakening, but this functional damage is still caused by myocardial stunning or myocardial cell death. A model that is more similar to human myocardial infarction - local hypoxia remains to be established on zebrafish. Together, these studies provide abundant evidence for the extraordinary cardiac regeneration of zebrafish. Figure 2.1 summarizes the types of injury models used in zebrafish regeneration studies.

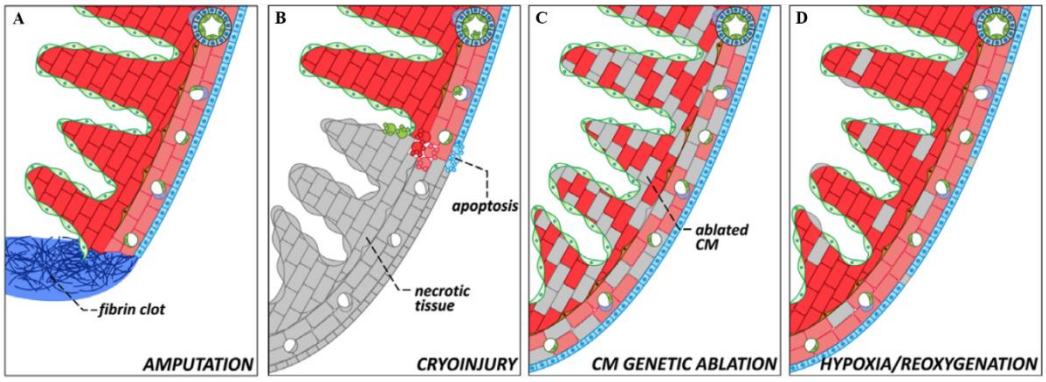


Figure 2.1 Four types of heart injury models commonly found in zebrafish heart regeneration studies [Gon17]

Figure A shows the surgical resection method, which quickly forms a fibrous clot. Figure B shows the hypothermia method. Tissue is necrotic at low temperatures, and cells in the vicinity of the necrotic area begin to undergo apoptosis. Figure C shows the myocardial cell genetic damage method, which specifically induces myocardial cell damage by genetic means. Figure D shows the hypoxia-reoxygenation method, which can only cause the death of some cardiomyocytes without histologically visible damage.

2.6.3 Amphibian and mammalian heart regeneration

The mechanism of cardiac maturation and regeneration has been studied for a long time and is highly controversial. During the study of this problem, from marine invertebrates to mammals ([WuJ96], [Tia15]), heart stab wounds, cuts, frostbite, injections of toxins, infections and infarctions have all been applied, but this problem remains a huge challenge. For most researchers, it is believed that the progressive maturation of the heart is accompanied by the loss of cardiac regenerative capacity. There are still many problems in this research which should be continued to explore, heart failure is a new public health problem, and myocardial cell deficiency is its main cause. There are 200-400 million cardiomyocytes in the human left ventricle, and myocardial infarction can eliminate 25 % of these myocardium cells in just a few hours [Mae14]. Recently, the only clinical solution is heart transplantation, however, it entails, problems regarding donors and the expensive surgery. Many scientists now focus on stem cell therapy, using human tissue-derived pluripotent stem cells for directional

differentiation. In addition, there is a problem of an abnormal heart rate after transplantation. Scientists have used a variety of methods including factor and small molecule regulation, and the effect is not so significant.

Unlike humans, many amphibians and fish are prone to regenerate limbs, appendages and internal organs after injury. Studies have shown that the zebrafish heart can be completely regenerated after it has been removed at the tip of the heart ([Pos02], [Lep06]). This large wound is effectively sealed by the initial fibrin clot, which is gradually replaced by degenerative heart tissue rather than by scar tissue ([Lep06], [Kik10], [Jop10], [Soo98]). This regenerative response involves a large amount of myocardial proliferation, and zebrafish cardiomyocytes show a higher degree of cell cycle activity than equivalent mammalian cells. Previous studies have shown that undifferentiated cardiac progenitor cells are the main source of zebrafish regenerative cardiac muscle cells [Kik10], but recent research shows that this main source is formed by pre-existing cardiomyocytes ([Jop10], [Soo98]). Many teams successively verified this method in different ways. In order to determine whether the source of new proliferation is a regenerative cardiomyocyte, two different zebrafish strains were constructed. In one transgenic line, all myocardium cells and their offspring express green fluorescent proteins. In contrast, myocardium progenitor cells maintain the expression of a red fluorescent protein. The experiment shows that in the top of the heart after excision, all the myocardium expressed the green fluorescent protein, indicating that the cardiomyocytes produced in the zebrafish heart regeneration are derived from

previous cardiomyocyte, other teams also used the same experimental design to draw the same conclusion [Soo98]. A similar approach also studies the mechanism of cardiac regeneration in mammals [Ler15], and as a result, double-transgenic mice are produced. Most of the above studies focus on the proliferation of existing cardiac muscle cells and difficult to be used to detect cardiac muscle cells formed by progenitor cells. In order to determine whether progenitor cells contribute to cardiomyocyte turnover, researchers performed genetic fate tracking experiments in transgenic mice [Ric17]. This system allows authors to distinguish between cardiac cells from pre-existing (fluorescently labelled) cardiac cells and cardiac muscle cells from unlabeled progenitor cells. They found that this progenitor cell did not contribute significantly during normal aging; however, the number of infarct-labeled cardiomyocytes decreased. When combined with the discovery that cardiomyocyte colonization rates are very low in normal and injured rodent hearts, these data suggest that the limited endogenous repair mechanisms in adult mammalian hearts differ from those in zebrafish and are more dependent on cardiomyocyte proliferation as opposed to cardiac-derived progenitor cells. The studies have shown that the zebrafish-like regeneration mechanism is a different mechanism from mammalian heart regeneration.

2.7 Research Gap

In metabolic flux analysis, differential equations could be directly established to represent the equilibrium of the reaction rate at steady state when trying to analyze

a small-scale model. In addition, one equilibrium equation could also be established for each possible isotope isomer. The relationships could be demonstrated between the reactions in the system by solving these equations, meanwhile detecting the latent reaction in which the reactants and products share the same molecular formula but arrange in a different order. However, the complexity increase as the system grows exponentially. Therefore, only small-scale models can be analyzed with the limited computing resources. In order to analyze similar metabolic processes from a more macro perspective, the differential equation model was not established in detail for each reaction and potential reaction but allowed a random error in measuring the ^{13}C isotope concentration of each reactant. The Markov Chain-Monte Carlo (MCMC) method is used to sample the random error 1000 times, and the reaction rate constant is calculated according to the sampled steady state concentration of each reactant, so that the reaction model can obtain a more accurate solution at the molecular level. However, this approach has difficulties to revealing the reaction mechanism inside the metabolic process. The potential of the ^{13}C experiment difficult to be fully exploited.

The traditional MFA method is to establish differential equations and solve them after the system reaches steady state. This kind of method is more effective when dealing with small and medium-sized reaction systems. But when modeling larger systems, or when some of the reactions in the system are sensitive to changes in experimental conditions, more complex equations needs to be created. At the same

time, when using the ^{13}C experiment, multiple isotopic isoforms of the same compound will increase the complexity of the equation. In metabolic disease-related models, such as insulin resistance processes and cardiac regeneration systems, multi-angle dynamic models can be combined with existing methods to better explore the flux analysis of metabolic processes and the in-depth study of these models. Therefore, a new dynamic model, which describes the forward process of the system from the perspective of transition probability, is proposed and it is expected to be effectively applied to each biological system model.

Chapter 3 Continuous-time Markov chain based flux analysis in metabolism

3.1 Introduction

The metabolic network can be analyzed from various angles, and MFA (metabolic fluxes analysis) is the most popular one ([Ste98], [Bai91]).

MFA is in fact a computational model of the intracellular metabolic network [Fis04]. The idea of MFA is to build a model containing flows and metabolites. Considering that the inflow is equal to the outflow when the metabolic process reaches its steady state, differential equations can be established to analyze the concentration of each metabolite at the steady state [Sel04]. In addition, the minimum error between analytical data obtained from the model and actual measurement data at the steady state can be considered as an indicator for assessment. MFA can be simulated by Monte Carlo method ([Sch12], [Kad06]), Bayesian estimation simulation [Jay08], and Markov process [Zha13a], or a hybrid method, like Markov Chain Monte Carlo (MCMC) [Kad06]. Analytical data can be obtained through computer programming [Cha11a] and then analyzed using Lagrange multiplier methods ([Wie97], [XuZ13]), and finally processed with MFA.

In order to figure out the reaction mechanism in the metabolic process, generally, this study uses the method of ^{13}C isotope labeling. From the perspective of experimental analysis，MFA is carried out according to the relative concentration of metabolites containing ^{13}C which can be obtained through 1H NMR (Hydrogen-

1 nuclear magnetic resonance) spectroscopy experiment. Theoretically, a model can be built to simulate the metabolic pathway of metabolites containing ^{13}C ([Sch12]; [Ben94]). The ^{13}C marking method can be expressed by Monte Carlo random simulation [Sch12]. A binary representation was used to express the reaction mechanism of metabolites containing ^{13}C by Selivanov and Puigjaner [Sel04]. Then, experimental and theoretical analysis can be compared to calculate the relative error which serves as the assessment indicator of the model [Gom01]. In this study, chemical dynamics, bioinformatics and mathematical statistics are combined to carry out MFA. Based on chemical dynamics, the simple mathematical relations between the instantaneous rate (v) of the reaction is used and the concentration of the reaction products (A and B) as follows:

$$v = k[A]^{\alpha}[B]^{\beta} \tag{3.1}$$

where, α and β are the coefficients of A and B in the chemical reaction, k is a constant indicating the chemical reaction rate which can be measured by NMR experiments ([Mil79], [Gre88]) or calculated by establishing correspondent differential Equations [Mic01]; Furthermore, the metabolic process can be seen as the transition process of carbon atoms which conform to the Markov process, i.e., no memory property [Mar55]. Therefore, a Markov model is established to carry out MFA (MMFA), and thus obtain the relative concentrations of various metabolites when the metabolic process reaches the steady state.

The traditional MFA method focuses more on the study of the state of steady state and reverses the reaction process with steady state as a result. But the dynamic

process deduced by the forward process can make the MFA method more flexible and can create more complex models that can describe biological processes. Therefore, the multi-angle dynamic model can also better simulate the isotope tracing experiments (such as the 13C labeling experiment), in order to reduce the various experimental costs consumed in the disease model research in order to better simulate the disease model. Unlike other methods, MMFA analyzes the metabolic model only through the transition probability. This approach is generic and could be applied to any metabolic system if all the reaction mechanisms in the system are known. For the purpose of verification, pentose phosphate pathway will be selected to calculate the relative concentrations of individual metabolites at the steady state. From the result, this study found out the equilibrium constants of some reversible reactions in the metabolic process and compared with the measured data.

3.2 Methods

Consider a simple metabolic system which contains m metabolites denoted as A_i (i =1, 2, …, m), N_i represents the number of carbon atoms in the carbon backbone of A_i. Firstly, suppose an initial metabolite A_1 is selected. In fact, any metabolite A_i (i =1, 2, …, m) can be selected. Besides, suppose A_1 has n carbon atoms, that is, $N_1 = n$ ($n \geq 1$). Our problem is to derive the relative concentration for each metabolite A_i (i =1, 2, …, m) compared to the initial concentration of A_1 when the entire metabolic system reaches its steady state. Assuming the initial concentration

of A_1 is defined as 1, while the initial concentration of A_i (i =1, 2, …, m) is defined as 0. The metabolic process can be regarded as a transfer process for carbon atoms existed in A_1, so the problem can be transformed to study the distribution of each metabolite when every carbon atom in A_1 reaches the steady state.

In order to solve the above problem, a serial number is assigned for each of the carbon atoms in A_1 referred as 1, 2,…,n. $X^h(t)$ represents the state of the hth carbon atom at time t, where the state space is $M = \{1, \dots, m\}$. For example, $X^h(t) = a$ ($a \in M$) indicates that the hth carbon atom of A_1 is in A_a at time t. $P_{ab}^h(t, t')$ stands for the transition probability for the hth carbon atom from A_a to A_b after time interval t', that is,

$$\mathrm{P_{ab}^h(t,t') = Pr\left(X^h(t+t') = b \middle| X^h(t) = a\right)} \tag{3.2}$$

The transition matrix for the hth carbon atom of A_i is expressed as:

$$\mathrm{P^h(t,t') = (P_{ab}^h(t,t'))_{m\times m}}$$

Thus, the density matrix Q can be obtained by:

$$Q^h(t) = \frac{dP^h(t,t')}{dt'} = \left(\frac{dP_{ab}^h(t,t')}{dt'}\right)_{m\times m} := (q_{ab}^h(t))_{m\times m} \tag{3.3}$$

If $a \neq b$, for the continuous-time Markov chain,

$$q_{ab}^h(t) = \frac{dP_{ab}^h(t,t')}{dt'}$$

$$= \lim_{t'\to 0} \frac{P_{ab}^h(t,t')}{t'}$$

$$= \frac{v_{a\to b}^{h}(t)}{[A_a(t)]} \tag{3.4}$$

where, $[A_a(t)]$ represents the relative concentration of A_a at time *t*.

If $a = b$, then

$$q_{ab}^{h}(t) = -\sum_{b=1, b\neq a}^{m} q_{ab}^{h}(t) \tag{3.5}$$

Assume that there are g_h reactions in the transfer process from A_a to A_b of the *h*th carbon atom. Then combine with (3.1),

$$\mathrm{v}_{\mathrm{a}\to\mathrm{b}}^{\mathrm{h}}(\mathrm{t}) = \sum_{\mathrm{w}=1}^{\mathrm{g_h}} k_w[A]^{\alpha}[B]^{\beta} \tag{3.6}$$

where, k_w represents the reaction rate constant of the *w*th (*w*=1,2, …, g_h) reaction in (3.1).

For other reactions without the transfer process of the *h*th carbon atom, then

$$v_{a\to b}^{h}(t) = 0 \tag{3.7}$$

All reactions are assumed to be non-complex here.

Assuming π_j^h is the probability that the *h*th carbon atom existed in A_i transfers to A_j when the system reaches the steady state. So the steady state equation is:

$$\begin{gathered} \left(\pi_1^h, \pi_2^h, \ldots, \pi_m^h\right) Q^h(\infty) = 0, \\ \text{and} \sum_{j=1}^{m} \pi_j^h = 1 \quad h = 1,2,\ldots,n \end{gathered} \tag{3.8}$$

The entire metabolic process in fact is the transition of carbon atoms existed in A_1, so that at the steady state, the total number of carbon atoms that have transferred to A_j is：

$$[A_1(0)]V_0(\pi_j^1 + \pi_j^2 + \cdots + \pi_j^n) \qquad j = 1,2,\dots,m \tag{3.9}$$

where, V_0 stands for the entire volume (assumed to be unchanged during the metabolic process). Since each metabolite A_j contains N_j carbon atoms, the total number of metabolites at the steady state is:

$$\frac{[A_1(0)]V_0\left(\pi_j^1 + \pi_j^2 + \cdots + \pi_j^n\right)}{N_j} \tag{3.10}$$

Therefore, the relative concentration of A_j at the steady state can be derived as:

$$\left[A_j(\infty)\right] = \frac{[A_1(0)]V_0(\pi_j^1 + \pi_j^2 + \cdots + \pi_j^n)}{V_0 N_j[A_1(0)]} = \frac{(\pi_j^1 + \pi_j^2 + \cdots + \pi_j^n)}{N_j} \tag{3.11}$$

By combining (3.4), (3.5), (3.6), (3.7), (3.8) and (3.11), a set of quadratic equations can be obtained. Then π_j^h which is again plugged into (3.11), is calculated so that the relative concentration for each metabolite compared with the initial concentration of A_i at the steady state can be obtained.

3.3 Application

In this section, the proposed MMFA approach is applied to flux distribution analysis in the pentose phosphate pathway.

Firstly, a schematic diagram of the pentose phosphate pathway is illustrated in Figure 3.1.

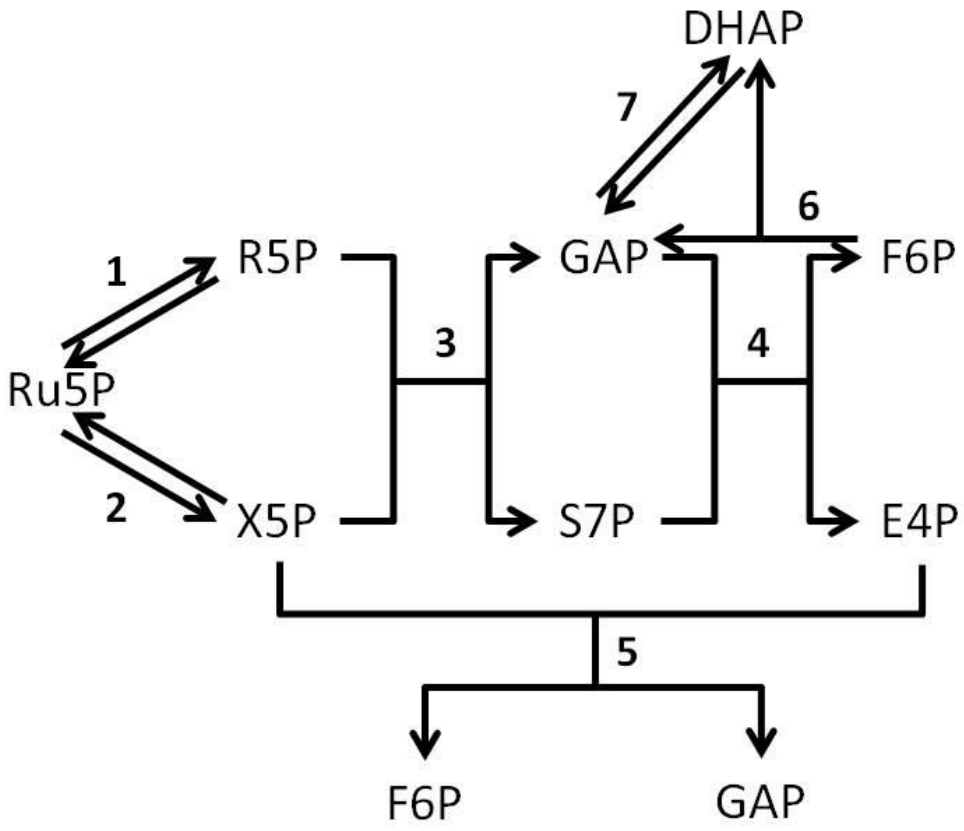


Figure 3.1 The process of the pentose phosphate pathway simulated in the model. Reaction 1 is Ru5P-R5P; reaction 2 is Ru5P-X5P, reaction 3 is R5P+X5P-S7P+GAP, reaction 4 is S7P-DHAP+E4P; reaction 5 is X5P+E4P-GAP+F6P; Reaction 6 is F6P-DHAP+GAP; Reaction 7 is DHAP-GAP. Abbreviation: F6P, fructose 6-phosphate; E4P,erythrose 4-phosphate; X5P, xylulose 5-phosphate; GAP, glyceraldehyde 3-phospate; DHAP, dihydroacetone phosphate; R5P, ribose 5-phosphate; Ru5P,ribulose 5-phosphate; S7P, sedoheptulose 7-phosphate.Details of the reaction are shown in reaction mechanism.

From this diagram, it is easily discovered that the entire system contains 8 metabolites and 7 reactions. The 8 metabolites are represented as A_1(Ru5P),

A_2(R5P), A_3(X5P), A_4(S7P), A_5(GAP), A_6(F6p), A_7(E4P), A_8(DHAP). The mechanisms of these 7 reactions are listed in Table 3.1:

Table 3.1 The mechanisms of these 7 reactions

1.		abcde	$\underset{k2}{\overset{k1}{\rightleftharpoons}}$	abcde				
		Ru5P		R5P				
2.		abcde	$\underset{k4}{\overset{k3}{\rightleftharpoons}}$	abcde				
		Ru5P		X5P				
3.	abcde	+ ABCDE	$\underset{k6}{\overset{k5}{\rightleftharpoons}}$	ABabcde	+	CDE		
	R5P	X5P		S7P		GAP		
4.		abcdefg	$\underset{k8}{\overset{k7}{\rightleftharpoons}}$	abc	+	defg		
		S7P		DHAP		E4P		
5.	abcde	+ ABCD	$\underset{k10}{\overset{k9}{\rightleftharpoons}}$	cde	+	abABCD		
	X5P	E4P		GAP		F6P		
6.		abcdef	$\underset{k12}{\overset{k11}{\rightleftharpoons}}$	abc	+	def		
		F6P		DHAP		GAP		
7.		abc	$\underset{k14}{\overset{k13}{\rightleftharpoons}}$	abc				
		DHAP		GAP				

In Table 3.1, k_w is a constant indicating the chemical reaction rate, shown in Table 2 [Mil79].

Table 3.2 Exact value of k_w

k_w	Exact value

k1	11.52
k2	26.72
k3	9.67
k4	6.91
k5	5.26
k6	4.46
k7	0.0017
k8	22.85
k9	267.78
k10	21.27
k11	0.000280
k12	1.83
k13	0.066
k14	1.38

In this pathway, by the analysis of the carbon atom's reaction mechanism in the entire process, it is found that the 1st carbon atom (that is, 'a' in Table 3.1) in A_1 can transfer back to any position (that is, 'a, b, c, d, e') in A_1. So the transfer process of each carbon atom is equivalent, this means $Q^1 = Q^2 = Q^3 = Q^4 = Q^5$. Then this study calculates the value of Q^1. After simulating the transfer process of the 1st carbon atom, it is hard to obtain the accurate probability of which product contains the 1st carbon atom if there are two or more products in the reaction. In

fact, in the entire pathway, the rate of each reaction is different. So in this study, the probability by the ratio of products' factors in the reaction is set as 1:1 in Table 3.1 obviously.

To test the effect of the MMFA approach, the relative concentration should be evaluated through the reaction equilibrium constant (*K*) in results section.

The model is calculated in detail as follows. The key point is to calculate the Q-matrix, combining with the equations (3.4) (3.5) (3.6) and (3.7) in Section 3.2, $q_{ab}^{i}(\infty)$ can be obtained as following:

$$q_{ab}^{i}(\infty) = \begin{cases} \dfrac{v_{a\to b}^{i}(\infty)}{[A_a(\infty)]} & a \neq b \\ -\sum_{b=1,b\neq a}^{m} q_{ab}^{i}(\infty) & a = b \end{cases}$$

For example, the reactions containing the transfer processes (from A_1 to A_2) of the 1st carbon atom existed in A_1 is:

$$A_1 \overset{k_1}{\rightarrow} A_2$$

So

$$v_{1\to 2}^{1}(\infty) = k_1[A_1(\infty)]$$

$$q_{12}^{1}(\infty) = \frac{v_{1\to 2}^{1}(\infty)}{[A_1(\infty)]} = k_1$$

As the same way, the 1st line of Q^1 (∞) is:

$$q_{12}^{1}(\infty) = k_1, q_{13}^{1}(\infty) = k_3, q_{14}^{1}(\infty) = \cdots = q_{18}^{1}(\infty) = 0,$$

$$q_{11}^{1}(\infty) = -\sum_{b=2}^{8} q_{1b}^{1}(\infty) = -k_1 - k_3$$

So,

$q_{21}^{1}(\infty) = k_2,$

$q_{22}^{1}(\infty) = -k_2 - k_5[A_3(\infty)]u,$

$q_{24}^{1}(\infty) = q_{25}^{1}(\infty) = 0.5k_5[A_3(\infty)]u,$

$q_{31}^{1}(\infty) = k_4,$

$q_{33}^{1}(\infty) = -k_4 - k_5[A_2(\infty)]u - k_9[A_7(\infty)]u,$

$q_{34}^{1}(\infty) = 0.5k_5[A_2(\infty)]u,$

$q_{35}^{1}(\infty) = 0.5k_5[A_2(\infty)]u + 0.5k_9[A_7(\infty)]u,$

$q_{36}^{1}(\infty) = 0.5k_9[A_7(\infty)]u,$

$q_{42}^{1}(\infty) = q_{43}^{1}(\infty) = 0.5k_6[A_5(\infty)]u,$

$q_{44}^{1}(\infty) = -k_7 - k_6[A_5(\infty)]u,$

$q_{47}^{1}(\infty) = q_{48}^{1}(\infty) = 0.5k_7,$

$q_{52}^{1}(\infty) = 0.5k_6[A_4(\infty)]u,$

$q_{53}^{1}(\infty) = 0.5k_6[A_4(\infty)]u + 0.5k_{10}[A_6(\infty)]u,$

$q_{55}^{1}(\infty) = -k_6[A_4(\infty)]u - k_{10}[A_6(\infty)]u - k_{12}[A_8(\infty)]u - k_{14},$

$q_{56}^{1}(\infty) = k_{12}[A_8(\infty)]u,$

$q_{57}^{1}(\infty) = 0.5k_{10}[A_6(\infty)]u,$

$q_{58}^{1}(\infty) = k_{14},$

$q_{63}^{1}(\infty) = q_{67}^{1}(\infty) = 0.5k_{10}[A_5(\infty)]u,$

$q_{65}^{1}(\infty) = q_{68}^{1}(\infty) = 0.5k_{11},$

$q_{66}^{1}(\infty) = -k_{10}[A_5(\infty)]u - k_{11},$

$q_{74}^{1}(\infty) = k_8[A_8(\infty)]u,$

$q_{75}^{1}(\infty) = q_{76}^{1}(\infty) = 0.5k_9[A_3(\infty)]u,$

$q_{77}^{1}(\infty) = -k_8[A_8(\infty)]u - k_9[A_3(\infty)]u,$

$q_{84}^{1}(\infty) = k_8[A_7(\infty)]u,$

$q_{85}^{1}(\infty) = k_{13},$

$q_{86}^{1}(\infty) = k_{12}[A_5(\infty)]u,$

$q_{88}^{1}(\infty) = -k_8[A_7(\infty)]u - k_{13} - k_{12}[A_5(\infty)]u,$

The rest entries of the matrix are 0.

Mentioned in the study, $Q^1 = Q^2 = Q^3 = Q^4 = Q^5$, so $\pi_j^1 = \pi_j^2 = \pi_j^3 = \pi_j^4 = \pi_j^5$

(j=1, 2,…, 8)

By equation (3.11), $[A_j(\infty)]$ is computed as following

For example,

$$[A_1(\infty)] = \frac{(\pi_1^1 + \pi_1^2 + \cdots + \pi_1^5)}{N_1} = \frac{5\pi_1^1}{N_1} = \frac{5\pi_1^1}{5} = \pi_1^1$$

Similarly,

$$[A_2(\infty)] = \pi_2^1,$$

$$[A_3(\infty)] = \pi_3^1,$$

$$[A_4(\infty)] = \frac{5}{7}\pi_4^1,$$

$$[A_5(\infty)] = \frac{5}{3}\pi_5^1,$$

$$[A_6(\infty)] = \frac{5}{6}\pi_6^1,$$

$$[A_7(\infty)] = \frac{5}{4}\pi_7^1,$$

$$[A_8(\infty)] = \frac{5}{3}\pi_8^1,$$

Put above 8 equations into the matrix Q, and combine the matrix with equation (3.8), π_j^h is computed by the MATLAB. Putting π_j^h into equation (3.11), the relative concentrate $[A_j(\infty)]$ is obtained. Then K with the relative concentrates is found.

3.4 Results

Using MATLAB, initial concentration of A_1 is inputted and a group of probabilities are obtained when the system reached the steady state. Also, the related concentration of each metabolism is obtained at the steady state. To test the approach, the reaction equilibrium constant *K* is selected to be the assessment indicator due to each reaction in the pentose phosphate pathway is a reversible one. The pentose phosphate pathway has been studied by many researchers ([Luo07], [Kle06], [Zha08]). Although it is simple and small, the most reactants and products in the pentose phosphate pathway are related with other metabolic systems. In this research, the 2nd, 6th and 7th reactions are selected to test the approach. Each *K* of these 3 reactions listed in Table 3.3.

Table 3.3 Each chemical reaction equilibrium constant *K* of the 2nd, 6th and 7th reactions

	Concentration ratio	Exact value	references
K1	[X5P]/[Ru5P]	1.4	Casazza and Veech (1986)

K2	[GAP]/[DHAP]	0.047	Connett et al. (1985)
K3	[GAP][DHAP]/[F6P]	0.086	Veech et al. (1969)

In MATLAB, the initial concentration (defined as *u*) is verified from 10*mmol/L* to 50*mmol/L* and the distance value is 5*mmol/L*, the results are shown in Tables 3.4, 3.5 and 3.6. *The running time is about 3-5 minutes for each solution.*

Table 3.4 The actual error and relative error between the K_1's estimation value and exact value when *u* changes from 10*mmol/L* to 50*mmol/L* with the distance value is 5*mmol/L*.

Value of *u* (*mol/L*)	Estimation value	Exact Value	Actual error	Relative error
0.01	1.399351	1.4	0.000649	0.0464%
0.015	1.399334	1.4	0.000666	0.0476%
0.02	1.39932	1.4	0.00068	0.0486%
0.025	1.399309	1.4	0.000691	0.0494%
0.03	1.3993	1.4	0.0007	0.0500%
0.035	1.399309	1.4	0.000691	0.0494%
0.04	1.399284	1.4	0.000716	0.0511%
0.045	1.399277	1.4	0.000723	0.0516%
0.05	1.399271	1.4	0.000729	0.0521%
average	1.3993061	1.4	0.0006939	0.0496%

Table 3.5 The actual error and relative error between the K_2's estimation value and exact value when *u* changes from 10*mmol/L* to 50*mmol/L* with the distance value is 5*mmol/L*.

Value of *u* (mol/L)	Estimation value	Exact Value	Actual error	Relative error
0.01	0.048174	0.047	0.001174	2.50%
0.015	0.048364	0.047	0.001364	2.90%
0.02	0.048555	0.047	0.001555	3.31%
0.025	0.048746	0.047	0.001746	3.71%
0.03	0.048937	0.047	0.001937	4.12%
0.035	0.048746	0.047	0.001746	3.71%
0.04	0.049321	0.047	0.002321	4.94%
0.045	0.049515	0.047	0.002515	5.35%
0.05	0.049709	0.047	0.002709	5.76%
average	0.048896	0.047	0.001896	4.03%

Table 3.6. The actual error and relative error between the K_3's estimation value and exact value when *u* changes from 10*mmol/L* to 50*mmol/L* with the distance value is 5*mmol/L*.

Value of *u* (mol/L)	Estimation value	Exact Value	Actual error	Relative error
0.01	0.097444	0.086	0.011444	13.31%

0.015	0.086965	0.086	0.000965	1.12%
0.02	0.081333	0.086	0.004667	5.43%
0.025	0.077802	0.086	0.008198	9.53%
0.03	0.075388	0.086	0.010612	12.34%
0.035	0.077802	0.086	0.008198	9.53%
0.04	0.072343	0.086	0.013657	15.88%
0.045	0.071343	0.086	0.014657	17.04%
0.05	0.070562	0.086	0.015438	17.95%
average	0.078998	0.086	0.009761	11.35%

The results are also shown in Figure 3.2, where solid lines represent the estimation values of K_1, K_2 and K_3, while dotted lines are the exact values. From Figure 3.2(a), the relative error of K_1 between the estimation value and the exact value is extremely small, so the estimation value is quite accurate. Also, the estimation value of K_2 is very close to its exact value, and its relative error is less than 6%, as seen in Table 3.5 and Figure 3.2(b). The relative error of K_3 is a little more than K_1 and K_2, however, the estimation value is very close to its exact value. The reason that the relative error of K_3 is a little bit large is that the corresponding metabolite F6P can also be produced by G6P in the entire metabolic system, but G6P is not included in the pentose phosphate pathway. In all, the results of this approach are quite reliable and accurate.

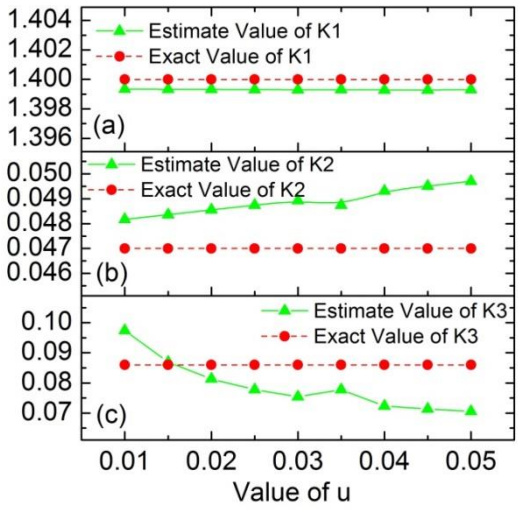


Figure 3.2 (a) is the estimation value of $K_1$when w changes from*10mmol/L* to *50mmol/L* with the distance value is *5mmol/L*. (b) is the estimation value of $K_2$when u changes from*10mmol/L* to *50mmol/L* with the distance value is *5mmol/L*. (c) is the estimation value of K_3 when *u* changes from*10mmol/L* to *50mmol/L* with the distance value is *5mmol/L*.

3.5 Conclusion

Metabolic Flux Analysis (MFA) is a powerful technique for determining intracellular metabolic fluxes in living cells. MFA has been found its widespread use in metabolic engineering, systems biology, biomedical research, and biotechnology.

In this Section, a brand-new approach, which combines Metabolic Flux Analysis and Continuous-time Markov Chain, has been put forward to analyze metabolic flux in the metabolic system. Based on the study of the pentose phosphate pathway discussed in the application section, this approach calculated the steady-state

concentration by the distribution of each carbon atom. The reaction equilibrium constant is obtained for analysis and comparison, and fthis approach is feasible and the results are accurate. The approach proposed in this study is generic, which can be applied to other metabolic systems.

Chapter 4 Continuous-time Markov chain with Monte Carlo based flux analysis in central metabolism

4.1 Introduction

To further extend the metabolic process in Figure 3.1, this section shows a more compressive central metabolic processes as Figure 4.1. The reaction has been increased from 7 reactions in Section 3 to 23 reactions in this section.

Metabolic networks can be analyzed from different perspectives, of which MFA (metabolic flux analysis) is the most commonly used ([Ste98], [Bai91]). MFA builds a model containing flow and metabolite. When the metabolic process reaches steady state, the inflow is equal to the outflow, and a differential equation can be established to analyze the concentration of each metabolite in steady state [Sel04]. The minimum error between the analytical data obtained from the model and the actual measured data at steady state can be considered as an indicator of the assessment. The MFA method can be simulated by the Monte Carlo method ([Sch12], [Kad06], Bayesian estimation simulation [Jay08], Markov process [Zha13a] and continuous time Markov process in Chapter 3. The integrated approach, MCMC, is also used to in this simulation [Kad06]. Data analysis can be obtained by computer programming [Cha11a], and then using the Lagrangian multiplier method for the analysis ([Wie97], [XuZ13]) of MFA.

In order to clarify the reaction mechanism in the metabolic process, the ^{13}C isotope labeling method is used. From the experimental analysis point of view, MFA was

carried out according to the relative concentration of the metabolite containing ^{13}C, which can be obtained by 1^{H}-NMR (hydrogen-1 nuclear magnetic resonance) spectroscopy. Theoretically, models could be built to simulate metabolic pathways containing ^{13}C metabolites ([Sch12], [Ben94]). ^{13}C labeling method can be represented by the Monte Carlo stochastic simulation [Sch12]. A binary representation can be used to represent the reaction mechanism for expressing metabolites containing ^{13}C by Selivanov and Puigjaner [Sel04]. Then, the experimental and theoretical analysis can be compared in order to calculate the relative error, which is used as the evaluation index of the model [Gom01].

In this study, the combination of chemical kinetics, bioinformatics and mathematical statistics is used to achieve MFA. Based on chemical kinetics, a simple mathematical relationship between the instantaneous rate of reaction (v) and the concentration of reaction products (A and B) is used as follows：

$$v = k[A]^{\alpha}[B]^{\beta} \tag{4.1}$$

where k is a constant indicating the rate of chemical reaction that can be measured by NMR experiments ([Mil79], [Gre88]), and calculated by establishing corresponding differential equations. In addition, the metabolic process can be seen as a transition process of carbon atoms consistent with the Markov process, which has no memory properties [Mar55]. This way, a Markov model for MFA （MMFA）is built, and the relative concentrations of various metabolites are obtained when the metabolic process reaches a steady state.

Not all values of the reaction rate constant k can be found in the literature and these values are critical to the calculation of the equation. To our knowledge, no studies have been performed to determine these missing *k* values, so it is decided to simulate the missing *k* values using the Monte Carlo method. The Monte Carlo method is a statistical simulation method that simulates an unknown distribution by generating a large number of random samples, and has been widely used in economics, biology, physics, and other fields with particle models. When a problem is so complex that an analytical solution difficult to be obtained, the Monte Carlo method can be used to effectively obtain a sufficiently accurate numerical solution. Schellenberger et al. [Sch12] simulated a large amount of ^{13}C-containing substances participating in the reaction in the Monte Carlo method and calculated the flow direction and distribution of labeled carbon atoms. Kadirkamanathan et al. [Kad06] also used the Monte Carlo method to simulate a Markov model of a glucose metabolism reaction process and calculate its stationary state. In this study, the Monte Carlo method is used to randomly generate the missing reaction rate constants *k*, and then calculate the corresponding reaction state, that is, the relative concentration of key reactants. According to the principle of the Monte Carlo method, when the sample of *k* is large enough, the expected concentration value obtained will be a valid estimate of the true relative concentration value in the glucose metabolism reaction.

4.2 Methods

Consider a simple metabolic system which contains *m* metabolites denoted as A_i (i =1, 2, …, *m*), N_i represents the number of carbon atoms in the carbon backbone of A_i.

Assuming that π_j^h is the probability that the *h*th carbon atom presented in A_i transfers to A_j when the system reaches the steady state. So the steady state equation similar to the chapter 3 is:

$$\begin{gathered}\left(\pi_1^h, \pi_2^h, \ldots, \pi_m^h\right) Q^h(\infty) = 0, \\ \text{and} \sum_{j=1}^{m} \pi_j^h = 1 \quad h = 1,2, \ldots, n\end{gathered} \tag{4.2}$$

Here, *Q* is the density matrix.

The relative concentration of A_j at the steady state can be derived as:

$$\left[A_j(\infty)\right] = \frac{(\pi_j^1 + \pi_j^2 + \cdots + \pi_j^n)}{N_j} \tag{4.3}$$

Then this study calculates π_j^h which is again plugged into (4.3), so that the relative concentration for each metabolite compared to the initial concentration of A_i at the steady state can be obtained. The reaction rate constants *k* can be randomly selected by the Monte Carlo method. A random sample is generated first. Then the necessary transformations or calculations on the sample based on the objective is carried out and the calculation result of the sample satisfies the acceptance condition is determined. Finally, the previous steps are repeated in large numbers

until the number of repetitions or the accuracy of the results meets the given conditions.

Therefore, randomly generate many reaction constants *k* and substitute them into the dynamic equation to solve it. Since the solution to this equation represents the concentration of each reactant in the sugar metabolism system, it should have a non-negative real solution, used a condition for accepting the set *k* is used. Samples with negative or no real solutions are rejected. In the simulation, a probability density based on the existing 14 reaction constants is established. Considering the two characteristics of known constants: 1 - non-negative real number, 2 - large difference in magnitude; it is assumed that the logarithm of the normal distribution is normal, and the eigenvalue of the normal distribution is the eigenvalue of the known sample. According to the Monte Carlo method, the simulation will end when the number of accepted samples reaches 100 sets and the mean of the relative concentrations is obtained. The simulation process will be repeated 10 times.

4.3 Application

The MMFA method is applied to the three processes of metabolism: pentose phosphate pathway, glycolysis, and TCA cycle. The specific flows of these three processes are shown in Figure 4.1 which it the extended process of Figure 3.1.

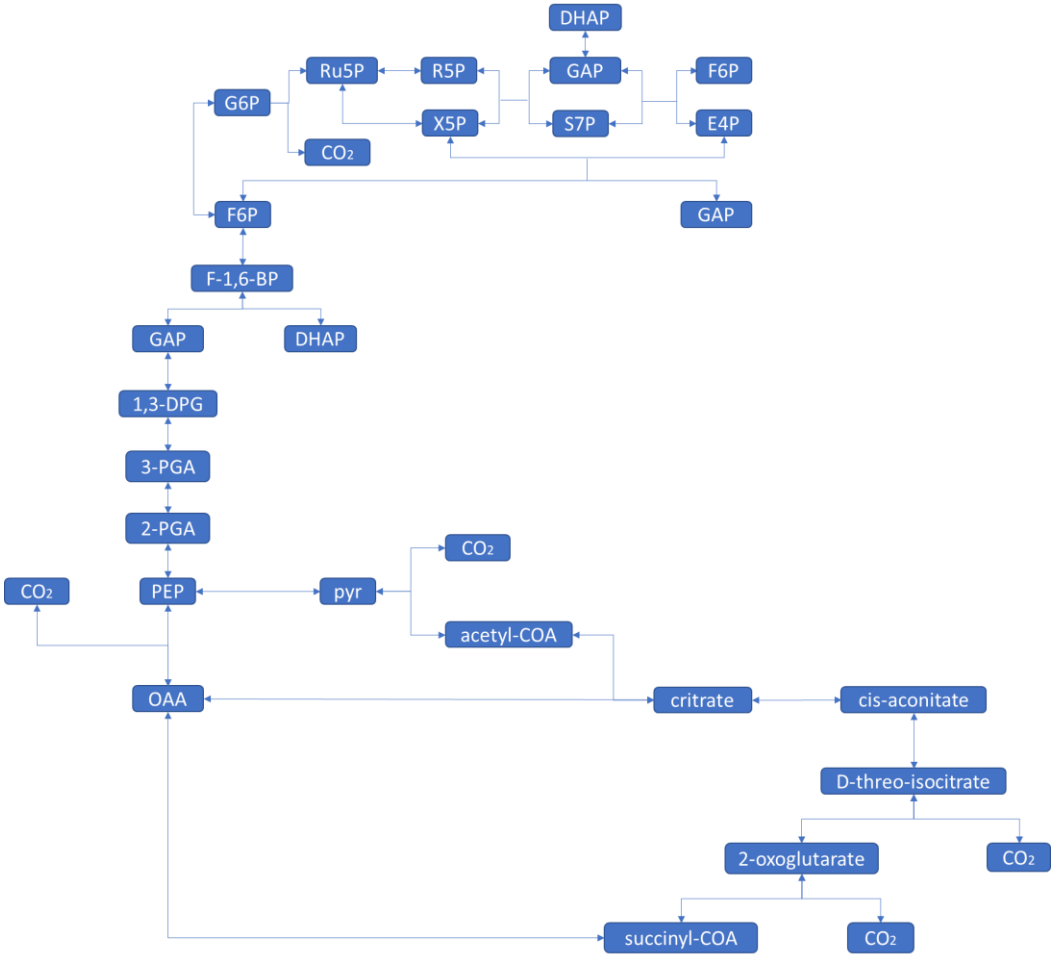


Figure 4.1 Central metabolic processes: pentose phosphate pathway, glycolysis, TCA cycle. [Abbreviations: G6P: Glucose 6-phosphate; F6P: Fructose 6-phosphate; F-1,6-BP: Fructose 1,6-bisphosphate; GAP: Glyceraldehyde 3-phosphate; DHAP: Dihydroxyacetone phosphate; 1,3-DPG: 1,3-Disphosphoglycerate; 3-PGA: 3-Phosphoglycerate; 2-PGA: 2-Phosphoglycerate; PEP: Phosphoenolpyruvate; OAA: Oxaloacetate; pyr: Pyrucate; acetyl-COA: acetyl coenzyme A; succinyl-COA: Succinyl-coenzyme A; Ru5P: Ribulose-5-phosphate; R5P: Ribose-5-phosphate; X5P: Xylulose 5-phosphate; S7P: Sedoheptulose 7-phosphate; E4P: erythrose 4-phosphate]

The whole process contains 23 compounds and 23 reactions, and the specific process and mechanism of each reaction will be shown in Table 4.1.

Table 4.1 The mechanisms of these 23 reactions

1.		abcdef	$\underset{k_2}{\overset{k_1}{\rightleftharpoons}}$	abcdef				
		G6P		F6P				
2.		abcdef	$\underset{k_4}{\overset{k_3}{\rightleftharpoons}}$	abcdef				
		F6P		F-1,6-BP				
3.		abcdef	$\underset{k_6}{\overset{k_5}{\rightleftharpoons}}$	abc	+	def		
		F-1,6-BP		GAP		DHAP		
4.		abc	$\underset{k_8}{\overset{k_7}{\rightleftharpoons}}$	abc				
		GAP		1,3-DPG				
5.		abc	$\underset{k_{10}}{\overset{k_9}{\rightleftharpoons}}$	abc				
		1,3-DPG		3-PGA				
6.		abc	$\underset{k_{12}}{\overset{k_{11}}{\rightleftharpoons}}$	abc				
		3-PGA		2-PGA				
7.		abc	$\underset{k_{14}}{\overset{k_{13}}{\rightleftharpoons}}$	abc				
		2-PGA		PEP				
8.		abcZ	$\underset{k_{16}}{\overset{k_{15}}{\rightleftharpoons}}$	abc	+	Z		
		OAA		PEP		CO_2		
9.		abc	$\underset{k_{18}}{\overset{k_{17}}{\rightleftharpoons}}$	abc				
		PEP		pyr				
10.		abc	$\underset{k_{20}}{\overset{k_{19}}{\rightleftharpoons}}$	ab	+	c		
		pyr		acetyl-COA		CO_2		
11.	abcZ	+	ab	$\underset{k_{22}}{\overset{k_{21}}{\rightleftharpoons}}$	aabbcZ			

	OAA		Acetyl-COA		citrate		
12.			aabbcZ citrate	$\underset{k_{24}}{\overset{k_{23}}{\rightleftharpoons}}$	aabbcZ cis-aconitate		
13.			aabbcZ cis-aconitate	$\underset{k_{26}}{\overset{k_{25}}{\rightleftharpoons}}$	aabbcZ D-threo-isocitrate		
14.			aabbcZ D-threo-isocitrate	$\underset{k_{28}}{\overset{k_{27}}{\rightleftharpoons}}$	aabcZ 2-oxoglutarate	+	b CO_2
15.			aabcZ 2-oxoglutarate	$\underset{k_{30}}{\overset{k_{29}}{\rightleftharpoons}}$	abcZ succinyl-COA	+	a CO_2
16.			abcZ succinyl-COA	$\underset{k_{32}}{\overset{k_{31}}{\rightleftharpoons}}$	abcZ OAA		
17.			abcdef G6P	$\underset{k_{34}}{\overset{k_{33}}{\rightleftharpoons}}$	abcde Ru5P	+	f CO_2
18.			abcde Ru5P	$\underset{k_{36}}{\overset{k_{35}}{\rightleftharpoons}}$	abcde R5P		
19.			abcde Ru5P	$\underset{k_{38}}{\overset{k_{37}}{\rightleftharpoons}}$	abcde X5P		
20.	abcde R5P	+	ABCDE X5P	$\underset{k_{40}}{\overset{k_{39}}{\rightleftharpoons}}$	CDE GAP	+	ABabcde S7P
21.			CDE GAP	$\underset{k_{42}}{\overset{k_{41}}{\rightleftharpoons}}$	CDE DHAP		
22.			abcdefg	$\underset{k_{44}}{\overset{k_{43}}{\rightleftharpoons}}$	abc	+	defg

			S7P			DHAP		E4P	
23.	abcde	+	ABCD	$\underset{k_{46}}{\overset{k_{45}}{\rightleftharpoons}}$		abABCD	+	cde	
	X5P		E4P			F6P		GAP	

In Table 4.1, k_w is a constant indicating the chemical reaction rate; the known k_w is shown in Table 4.2 [Mil79].

Table 4.2 Exact value of k_w

k_w	Exact value
k5	0.000280
k6	1.83
k35	11.52
k36	26.72
k37	9.67
k38	6.91
k39	5.26
k40	4.46
k41	1.38
k42	0.066
k43	0.0017
k44	22.85
k45	267.78
k46	21.27

The remaining *k* in the reaction mechanism was simulated by the Monte Carlo method. According to the MMFA method, the relative concentration of the compound can be calculated. Throughout the process, by analyzing the transfer mechanism of carbon atoms, it is found that the first carbon atom in the first compound can be transferred back to any position in it.

4.4 Results

Using MATLAB, the initial concentration of G6P is entered. When the system reaches a steady state so, a set of probabilities is obtained. The relevant concentrations of each metabolism is obtained at steady state. In order to test this method, the reaction equilibrium constant *K* is set as an evaluation indicator because each reaction in the glucose metabolism pathway is reversible. In this study, the 20^{th} and 23^{th} reactions is selected to test the approach. The *K* values for these 2 reactions are listed in Table 4.3.

Table 4.3 Each chemical reaction equilibrium constant *K* of the 20^{th} and 23^{th} reactions

	Concentration ratio	Exact value	references
K1	[X5P]× [R5P]/([GAP]× [S7P])	0.48	Casazza and Veech (1986)

K2	[GAP] ×[F6P] /([E4P] × [X5P])	29.7	Casazza and Veech (1986)

In MATLAB, 100 *k*-values are randomly selected to calculate the mean of the relative concentrations and 10 groups are selected for error comparison. Table 4.1 shows the actual and relative errors between the K_1's estimation value and the exact value among 10 random groups and Table 4.5 shows the actual and relative errors between the K_2's estimation value and the exact value among 10 random groups. *The running time is about 40 hours for each group of 20000 solutions. The running time for 23 reactions model is longer than the 7 reactions model. Cause in 23 reactions model, we need to randomly select the reactions rate constants k first and about 100 effective solutions can be obtained from 20000 solutions.*

Table 4.4 The actual and relative errors between the K_1's estimation value and the exact value among 10 random groups.

Group order	Estimation value	Exact Value	Actual error	Relative error
1	0.4766	0.48	0.0034	0.70%
2	0.4787	0.48	0.0013	0.27%
3	0.4808	0.48	0.0008	0.17%
4	0.4812	0.48	0.0012	0.24%
5	0.4802	0.48	0.0004	0.04%

6	0.4806	0.48	0.0006	0.13%
7	0.4816	0.48	0.0016	0.33%
8	0.4806	0.48	0.0006	0.13%
9	0.4794	0.48	0.0006	0.13%
10	0.4835	0.48	0.0035	0.73%
average	0.4803	0.48	0.0003	0.0663%

Table 4.5 The actual and relative errors between the K_2's estimation value and the exact value among 10 random groups.

Group order	Estimation value	Exact Value	Actual error	Relative error
1	29.2707	29.7	0.4293	1.45%
2	27.1370	29.7	2.5630	8.63%
3	30.3595	29.7	0.6595	2.22%
4	31.3304	29.7	1.6304	5.49%
5	27.6955	29.7	2.0045	6.75%
6	32.0134	29.7	2.3134	7.79%
7	27.8238	29.7	1.8762	6.32%
8	29.2381	29.7	0.4619	1.56%
9	29.0440	29.7	0.6560	2.21%
10	30.5122	29.7	0.8122	2.73%
average	29.4425	29.7	0.2575	0.867%

The results are shown in Figure 4.2 and Figure 4.3, the circle represents the mean of the random array; and the line represents the exact value. It is found that the circles are oscillating around the straight line of the exact value. The relative error between the K_1 estimated value and the exact value is extremely small, so the estimation value is quite accurate. The estimation value of K_2 is a little higher than K_1, but also very close to its exact value, and its relative error is less than 10%. The reason for the relative error of K_2 being a little bigger is that the corresponding metabolite F6P is also part of fructose and mannose metabolism, which generate GDP sugar for further glycosylation. In all, it is concluded that the results of this approach are quite reliable and accurate.

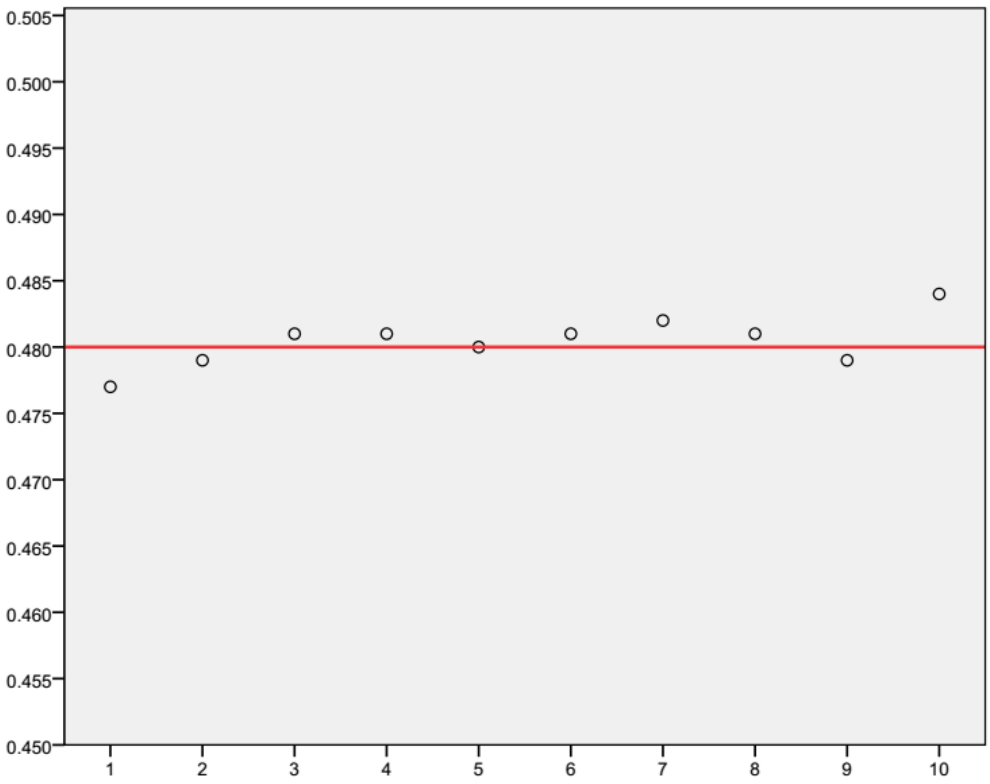


Figure 4.2 Results of K_1 among 10 random groups.

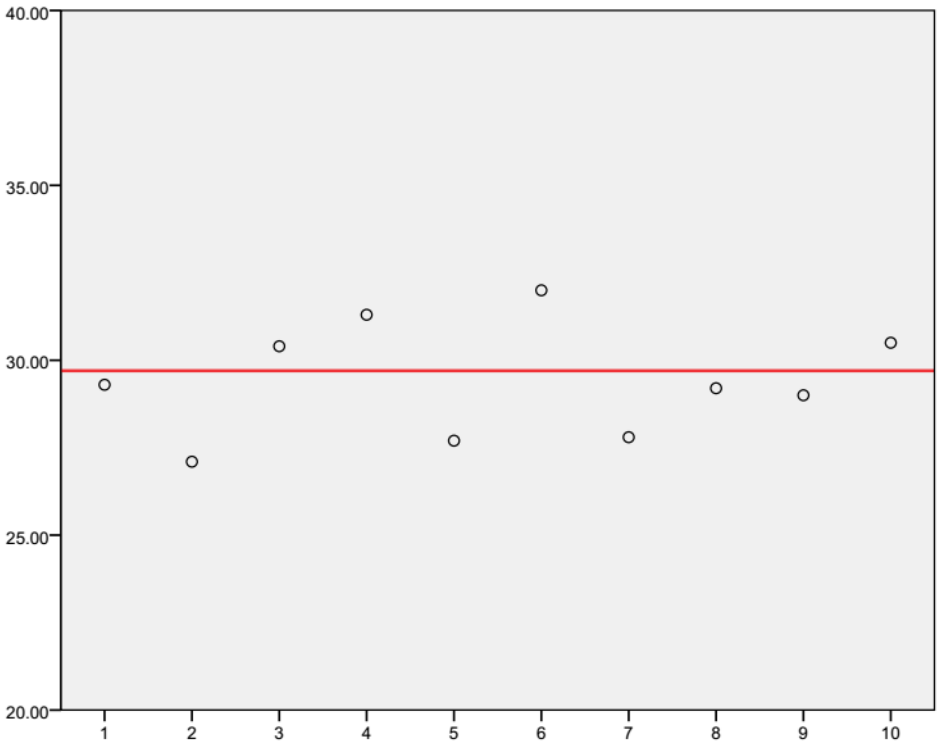


Figure 4.3 Results of K_2 among 10 random groups.

4.5 Conclusion

Metabolic flux analysis (MFA) is an effective method for determining metabolic fluxes in living cells. MFA is widely used in metabolic engineering, systems biology, biomedical research and biotechnology. In this study, the combination of the continuous time Markov chain and the Monte Carlo methods is applied to the three core processes in glucose metabolism. The Monte Carlo method is used to solve the problem of reaction rate unavailability and to calculate the equilibrium concentration with the distribution of each carbon atom when all the reactions reach steady state. The results show that the method is feasible; and the results are

accurate. The methods presented herein are generic and can be applied to any metabolic system. Central metabolism plays an important role in heart diseases. Abnormal metabolism leads to an accumulation of different metabolites and further causes cardiovascular diseases, especially under stresses like hypoxia or arrhythmia. The MFA model of central metabolism provides a great tool for prevention of metabolism-related cardiovascular diseases

Chapter 5 Multiple homeostasis in the insulin receptor reveals the cause of insulin resistance in Type 2 diabetes

5.1 Introduction

In Section 3 and Section 4, dynamic MFA model was applied to the central metabolic pathways. These metabolic processes are of great importance in some disease models, such as Type_2 diabetes and related disease models. Cause Type 1 diabetes is the congenital disease. The mechanism and model are totally different from the Type 2 diabetes. This study mainly focused on the Type 2 diabetes. In order to apply the model more extensively and in depth, it is experimental to apply the model in metabolic-related disease models. Insulin signaling pathway is selected to simulate. However, it is found that the phenomenon of multiple homeostasis exists in the process. In this section, multiple homeostasis in the insulin signaling pathway is discussed.

Insulin resistant (IR) is one of the origins in Type II Diabetes. But its regulation mechanism has not been figured out. As one of the protein tyrosine phosphatase (PTP), LAR (leukocyte antigen-related) can restrain the signal transduction of insulin. In the insulin resistance model of animals or human beings, the expression quantity of LAR dynamically increased [Lev50]. By Single-strand conformation polymorphism (SSCP) and automatic DNA sequencing [Fre71], the whole sequence of LAR gene and whether there some specific gene mutations lead the insulin to restrain can be found [Sor90]. 276b base pairs on the upstream region of transcription initiation site (TSS) have promoter activity. And among 589 non-

diabetic patients in the middle East Coast area of Italy, the mutation of point T→A on -127 (mutation rate: 5%) is related to low BMI ([Acc96], [Acc99], [Tav93]), low waist circumference, low blood pressure and the lower rate of albumin-to-creatinine in urine. To further quantify the influence of mutation on the point to overweight, the group has been divided to 3 groups based on the sizes of BMI. The probability of the group with the mutation point to the biggest size BMI group is 60% lower than the smallest size BMI group [Tay98]. The other group with 307 respondents in Eastern Sicily, 13-point mutation carriers showed low triglyceride, and in OGTT (oral glucose tolerance testing) experiment [Acc96], they showed the high sensitive to Insulin. In one word, the point of mutation of LAR gene promoter area in two Caucasus groups shows the similar phenotype in the sensitivity to insulin [Bru97]. These findings indicate LAR gene may plan an important role in the sensitivity to insulin and other related disease models [Kid00]. Until now, many evidences indicate that the activation to the physiological processes by insulin needs the phosphorylation of the tyrosine kinase to insulin receptors ([Kul99], [Hub94]). The researches of biochemistry area show that receptor kinase can be regulated by internal Protein Tyrosine Phosphatase (PTP) [Ren08]. To improve transmembrane Protein Tyrosine Phosphatase, LAR can regulate the insulin receptor signal transduction process. Antisense RNA can be used to restrain the expression of LAR gene [Lou06]. The Hormone dependence autophosphorylation of insulin receptor raised about 150% after restraining LAR. At the same time, the activity of insulin receptor tyrosine kinase also increased

35%. Meanwhile [Che08], constraining LAR difficult to make influence on non-hormone dependence autophosphorylation [McK06], neither the basal phosphorylation of insulin receptor. What count most is that the restraining LAR makes the downstream hormone dependence phosphatidylinositol 3 kinase activity increased 350% [Nef03]. All the vivo tests data proved that LAR takes part in restraining process of insulin receptor signal transaction.

GRB2 protein has one SH2 and two SH3 modular structure models [War04]. And this is also a highly conservative regulation p21-Ras signal transaction pathway mechanism. Under the situation with the existing of insulin, GRB2 combines with two tyrosine phosphorylation proteins to a stable compound [Sla06]. One is the main insulin receptor substrate (IRS) IRS-1; the other is oncogene Shc with SH2 modular structure models [Whi03]. The interaction of GRB2 and two proteins need the Ligand activation of insulin receptor, and the interaction is formed by the direct combination of SH2 modular structure model on GRB2 with Tyrosine phosphorylation on Shc [Zic05]. Though GRB2 can combine with both proteins, the phosphorylation of GRB2 itself difficult to be tyrosine. In other word, GRB2 is not the direct substrate of insulin receptor. What's more, the research showed a short sequence motif (YV/IN) can directly combine with the SH2 modular structure model on GRB2 in IRS-2, EGFR and Shc [Rem05]. Interestingly, GRB2 and Phosphatidylinositol 3 (PI3) kinase can simultaneously combined to the different Tyrosine sites on IRS-1 molecule, which indicates that IRS-1 may become the hardcore of the big signal transduction complex. The combination of

IRS-2 [Epa06], EGFR and Shc may play the key role in the regulation of the p21-Ras pathway and downstream effectors [Thi08].

The multiple steady states had been found in catalytic combustion with considered as a monolith catalytic reactor [Jam02]. However, multiple homeostasis problems are not taken into consideration in the dynamic process of biology systems. In this study, various stable steady points will be found in the insulin signaling pathway. The stable steady points in this research are different with the equilibrium points in the traditional chemical reaction dynamic model. In traditional chemical reaction dynamic model, equilibrium points represent the state in which the reaction rate and the concentration change are both 0. Under the small disturbance, equilibrium points will change immediately. However, the stable steady points are also the minimum potential energy states. So, the system can return to the stable steady states under the disturbance.

5.2 Methods

Suppose there are N substances, X_1, X_2…, X_N, taking part in M reactions in the system. The concentration of each substance is recorded as $[X]_k$, k=1, 2…, N, and the state vector of the system is defined as $[\mathrm{X}] = [[\mathrm{X}]_1, [\mathrm{X}]_2, \cdots, [\mathrm{X}]_\mathrm{N}]^\mathrm{T} \in \mathbb{R}^\mathrm{N}$. The state change vector of the jth reaction is $\mathrm{v_j} \in \mathbb{Z}^\mathrm{n}$, the reaction rate is $\bar{\mathrm{a}}_\mathrm{j}(\mathrm{X}),\ \mathrm{j} = 1,2,\cdots,\mathrm{M}$. $\mathrm{v_{ij}}$ represents the number i component of $\mathrm{v_j}$, which shows the variation amount of number i reactant after number j reaction is done. To describe the chemical reactions perfectly, the model needs to describe the velocity of each

reaction underway. There are quite a lot of experimental laws depicting the velocity and the most important one is the mass action law. If one of the reactions is

$$c_1X_{n_1} + c_2X_{n_2} + \cdots + c_rX_{n_r} \to d_1X_{m_1} + d_2X_{m_2} + \cdots + d_rX_{m_s} \tag{5.1}$$

The state change vector of this reaction is $v_{jn_k} = -c_k,\ k = 1,2,\cdots,r,\quad v_{jm_k} = d_k,$ $k = 1,2,\cdots,s.$ The mass action law shows that the reaction velocity of this reaction is

$$\bar{a}([X]) = \tilde{k}[X]_{n_1}^{c_1}[X]_{n_2}^{c_2}\cdots[X]_{n_r}^{c_r}. \tag{5.2}$$

where, $\tilde{k}$ is a constant irrelevant to the state of the system called reaction velocity constant.

Another common reaction velocity law is the Michaelis-Menten law of enzymatic reaction. For an enzymatic reaction:

$$X_n + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C \to X_m + E \tag{5.3}$$

E represents the enzyme; *C* represents the compound of the enzyme and the substrate. The Michaelis-Menten law shows the velocity of this reaction is

$$\bar{a}([X]) = \frac{V_M[X]_n}{K_M + [X]_n} \tag{5.4}$$

In which $K_M = k_{-1}/k_1$ is called Michaelis constant, $V_M = k_2([E] + [C])$. If only focus on the journey from the substrate to the product $X_n \to X_m$, the velocity will be the one in (5.3). The constants $V_M,\ K_M$ are determined by experiments.

With these laws describing reaction velocity [Kee09], the evolving pattern of the concentration of the substance in system over time is:

$$\frac{d[X]}{dt} = \sum_{j=1}^{M} \bar{a}_j([X])v_j \tag{5.5}$$

More commonly, one ordinary differential equation can be used to describe the chemical reaction system:

$$\frac{d[X]}{dt} = b([X]),\ [X] = x_0 \tag{5.6}$$

It is a most typical research method, based on the deterministic dynamics model, to collect all kinds limit sets and analyze the stability of singular point and ramification while explaining the biological meaning corresponding to the mathematics concepts.

In practice, the chemical reactions in living bodies will inevitably be affected by noises. Noises can be divided into 2 types: one being from the outside surroundings called external noise. This type of noise is considered irrelevant to the state of the system hence a white noise can be used to describe it. The dynamics equation of the system is altered as

$$\frac{d[X]}{dt} = b([X]) + \sigma\dot{W}_t,\ [X] = x_0 \tag{5.7}$$

In which $\dot{W}_t$ is standard white Gaussian noise, $\sigma > 0$ features the intensity of the noise. (5.6) is a proper model for description of the chemical reaction with external noise.

The other one is the noise intrinsic of the biochemical reaction system, called

intrinsic noise. In recent years, more and more experiments indicate that biochemical system is one intrinsic random dynamic system. The well-known experiment [Elo02] has observed the intrinsic randomness on cell level. The author marked 2 gene points on the colibacillus's DNA molecule, both controlled by the same promoter and with equal distance from the promoter, one green and one yellow. It turns out that not only are the expression amounts of the 2-fluorescein perturbed by noises but also not synergetic. The not-synergetic fact shows that even without surrounding noises, the chemical reaction of genetic expression itself is random.

Intrinsic randomness is more significant when there are fewer molecules involved in the reaction, yet living bodies are often modulated by a few key molecules with small amount. For instance, there are usually only several or a couple more of one type enzyme molecules in the enzymatic reactions of the cell. The mRNA in genetic modulation shares the same range, with probably only one or two DNA molecules out there. The randomness of these key molecules will also affect vital process. On a subtler unimolecular level, in a living cell experiment conducted in lab, thanks to the fluorescein labeling technique, it is observed that the synthesis of protein showed significant randomness [LiG11].

When there are few molecules, to describe the intrinsic randomness, it is not good enough to do so with concentration and ordinary differential equation. The classic model for it is the Gillespie dynamics. The occurrence of chemical reaction is the result of collisions of molecules. The Gillespie dynamics assume that the collisions

are fully and frequently performed, yet the effective collisions that can cause chemical reactions are scarce. With the ignorance of uneven space as the assumption, chemical reactions are depicted as a process of Markov jump in continuous time and discrete state. It was initially used to simulate chemical reactions in biology [Gil77].

Different from the deterministic model, the number instead of the concentration of the molecules is applied for describing chemical reaction system. Suppose the number of molecules of N reactants are as $X_k(t) \in \mathbb{Z}, k = 1,2,\cdots,N$, and $X_t = [X_1, X_2, \cdots, X_N]^T \in \mathbb{Z}^N$ are the state vector of the system. The state change vector of number j reaction is v_j, $j = 1,2,\cdots,M$. To feature each reaction velocity is the reaction velocity function (or propensity function) $a_j(x): \mathbb{R}^N \to \mathbb{R}$, $j = 1,2,\cdots,M$. For a reaction, reaction velocity is

$$a(X) = k\binom{X_{n_1}}{c_1}\binom{X_{n_2}}{c_2}\cdots\binom{X_{n_r}}{c_r} \tag{5.8}$$

If $n < m$, then $\binom{n}{m} = 0$. (5.7) shows that reaction velocity is proportional to frequency of collisions. Here the constant of proportionality k is not necessarily the same with $\tilde{k}$ from the mass action law.

The occurrence of chemical reaction is however a process of Markov jump process. When its current state is X_t, each reaction needs to wait for an independent, normal distributed time $\tau_j \sim \exp(a_j(X_t))$ to happen. After the first reaction, the system state changes while remaining the same between two reactions. Applying the form of random integral equation, there is:

$$X_t = X_0 + \sum_{j=1}^{M} P_j\left(\int_0^t a_j(X_s)ds\right)v_j \tag{5.9}$$

In which $\{P_j(t)\}_{j=1,2,\ldots,s}$ is a row of independent Poisson process with velocity of 1. The chronological evolvement of $p(X,t)$ the probability distribution of X_t, meet with the so-called chemical master equation [Gar85].

$$\partial_t p(X,t) = \sum_{j=1}^{M}\left(a_j(X - v_j)p(X - v_j, t) - a_j(X)p(X,t)\right) \tag{5.10}$$

It is proposed to apply stochastic dynamics of chemical reaction approach to flux distribution analysis in insulin signaling pathway.

5.3 Application

First, the schematic diagram of the process is in Figure5.1. Arrows indicate upstream substance activates the downstream ones, while vertical lines indicate the opposite. INSR activates both GRB2 and IRS, and GRB2 can also be activated by IRS. LAR works as a global inhibitor with its phosphorylase activity. The relationships between these proteins are reversible due to the existence of feedback mechanism.

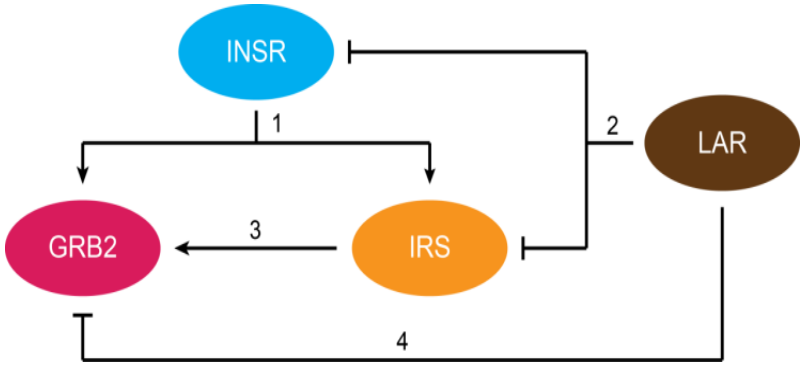


Figure 5.1 Process of insulin signaling pathway.

As for differential equations:

$$\frac{dx}{dt} = f(t, x), \quad x \in \mathbb{R}^n \tag{5.11}$$

If $f(t, \tilde{x}) = 0$ for all t, the point $\tilde{x}$ is an equilibrium point. The differential equations could be linearized into the below equation:

$$\frac{dx}{dt} = Ax \tag{5.12}$$

A is the Jacobian matrix at the point $\tilde{x}$ of the system. Equilibria can be classified by looking at the signs of the eigenvalues of the linearization of the equations about the equilibria. If all eigenvalues have negative real part, the equilibrium is a stable point. If at least one eigenvalue has negative real part and at least one has real part, the equilibrium is a saddle point. Also, all solutions are required to be positive real. The negative real solutions can be transformed to positive real solution by the translation.

Here shows the function $F(X) = (f_1, \cdots, f_n)(x)$ and the variable $X = (x_1, \cdots, x_n)$. As for ordinary differential equation $\partial_t X = F(X)$, require $F(X) = 0$, then obtain solution $x_0, x_1, \cdots, x_p$.

Here,

$$DF = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \tag{5.13}$$

Taking consideration of translation: $Y = X - A$. $A \pm (a_1, \cdots, a_n)$ is constant vector, the corresponding equation is $\partial_t Y = G(Y)$. $G(Y) = F(Y + A)$ and

$F(X) = 0 \Leftrightarrow G(Y) = 0$, then $DF(X) = DG(Y)$, so the eigenvalues of the equations are invariant after translation.

For the model, the reactants will be affected by external interference. Taking an example of IRS, some proteins such as ptp1b, jnk and ikk all have inhibitory effects on IRS. These inhibitory effects are expressed as constant B_0 in the model. Similarly, INSR, LAR and GRB2 are also inhibited by other proteins, expressed as A_0, C_0, D_0. So, in the model, new variables can be found, $A' = A - A_0, B' = B - B_0, C' = C - C_0, D' = D - D_0$.

From the above discussion, the model needs to find eigenvalue vectors in which four components are all negative and the corresponding solutions are all real.

The ordinary differential equations are as follow:

$$\begin{aligned}
\frac{dA}{dt} &= k_1'BC - k_1A - k_2D + a_0 \\
\frac{dB}{dt} &= k_1A - k_1'BC - k_3B + k_3'C - k_2D \\
\frac{dC}{dt} &= k_1A - k_1'BC + k_3B - k_3'C - k_4D \\
\frac{dD}{dt} &= k_2'AB + k_4'C
\end{aligned} \tag{5.14}$$

Here, *k* and k' are the equivalent reaction rate constants. And

$$\frac{dA}{dt} = \frac{dB}{dt} = \frac{dC}{dt} = \frac{dD}{dt} = 0 \tag{5.15}$$

Eigenvalue matrix M can be showed as follows and $|M| < 0$

$$M = \begin{pmatrix} -k_1 & -k_1'C & k_1'B & -k_2 \\ k_1 & -k_1'C - k_3 & k_3' & -k_2 \\ k_1 & -k_1'C + k_3 & -k_1'B - k_3' & -k_4 \\ k_2'B & k_2'A & k_4' & 0 \end{pmatrix} \tag{5.16}$$

5.4 Results

For specific insulin concentration, the activated initial concentration of INSR can be defined as $a_0 = 0.02$. According to the relation between the reactants and the orders of magnitude of the reaction rates [Mar85], the ranges of parameters are in the Table 5.1. Under the insulin stimulation, the insulin signaling pathway will mainly proceed to positive. So, there is a difference between the orders of magnitude of *k* and *k*'. As the reaction from LAR to GRB2 is an indirect reaction, the model difficult to confirm whether it is a positive feedback control or negative feedback control. The range of k_4' is defind as [-0.1, 0.1].

Table 5.1 Range of parameters

Parameters	Range
k_1	[0.1, 20]
k_1'	[0.01, 1]
k_2	[0.1, 20]
k_2'	[0.01, 1]
k_3	[0.1, 20]
k_3'	[0.01, 1]
k_4	[0.1, 20]
k_4'	[-0.1, 0.1]

As all the parameters are within the range, the discussion of the parameters will be based on the control variable method.

Firstly, it is crucial to discuss the single parameter traversal situations. Take traversal in range of any parameter; fix the rest eight parameters and analyze three corresponding eigenvalues of three sets of real solutions. Figure 5.2 shows three corresponding eigenvalues of three sets of real solutions are.

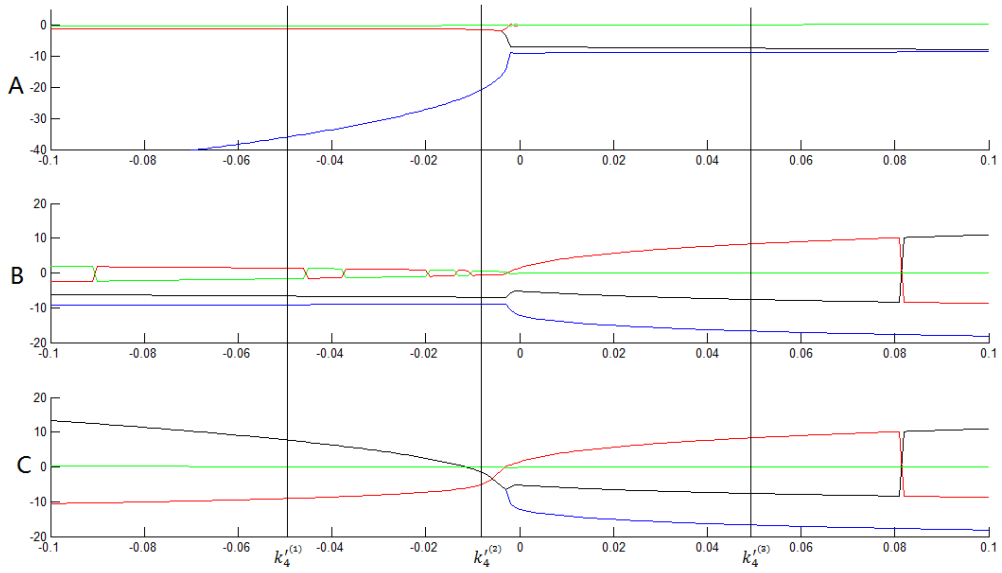


Figure 5.2 Three corresponding eigenvalues of three sets of real solutions are separately shown in A, B, C when k_4' take traversal in [-0.1, 0.1]. (The green line is the concentration of the INSR, the red line is the concentration of the LAR, the blue line is the concentration of the GRB2, the black line is the concentration of the IRS)

Selecting k_4' as the small order of magnitude of the parameters, the model take traversal in [-0.1, 0.1] and fix the rest parameters (Figure.5 2). As seen from Figure 5.2, three corresponding eigenvalues of three sets of real solutions are separately shown in A, B, C. The horizontal axis represents the value of k_4' and four curves represent the corresponding value of four components of each eigenvalue vector in subgraphs A, B, C. When $k_4' = k_4'^{(1)}$, only one set of eigenvalue vector can be obtained in which four components are all negative and the corresponding solutions are all real (A). So there is only one stable point. When $k_4' = k_4'^{(2)}$, two sets of eigenvalue vector can be obtained in which four components are all negative and the corresponding solutions are all real (A, C). So there are two stable

points. When $k_4' = k_4'^{(3)}$, none of set of eigenvalue vector can be obtained in which four components are all negative and the corresponding solutions are all real. As a result, there is no stable point.

Selecting k_2 as the large order of magnitude of the parameters, the model take traversal in [0, 20] and fix the rest parameters (Figure. 5.3). As seen from Figure 5.3, three corresponding eigenvalues of three sets of real solutions are separately shown in A, B, C. The horizontal axis represents the value of k_2 and four curves represent the corresponding value of four components of each eigenvalue vector in subgraphs A, B, C. When $k_2 = k_2^{(1)}$, none of set of eigenvalue vector can be obtianed in which four components are all negative and the corresponding solutions are all real. So there is no stable point. When $k_2 = k_2^{(2)}$, only one set of eigenvalue vector can be obtained in which four components are all negative and the corresponding solutions are all real (A). So there is only one stable point. When $k_2 = k_2^{(3)}$, two sets of eigenvalue vector can be obtained in which four components are all negative and the corresponding solutions are all real (A, C). So there are two stable points.

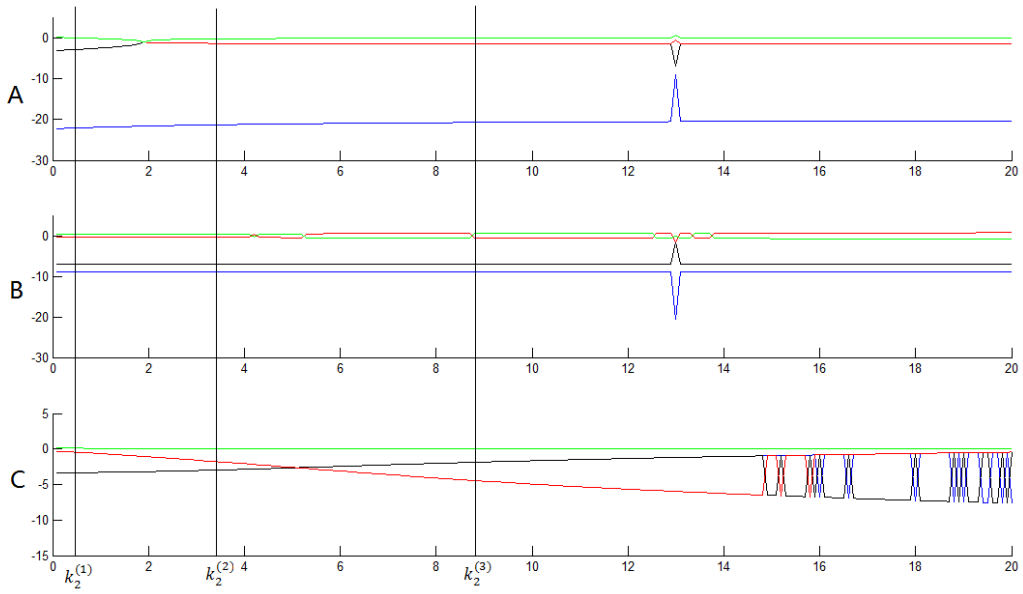


Figure 5.3 Three corresponding eigenvalues of three sets of real solutions are separately shown in A, B, C when k_2 take traversal in [0, 20]. (The green line is the concentration of the INSR, the red line is the concentration of the LAR, the blue line is the concentration of the GRB2, the black line is the concentration of the IRS)

The single parameter traversal analysis of other parameters is similar with the above two discussions in Figure 5.2 and Figure 5.3. Detailed data can be found in the appendix II.

Secondly, two parameters were discussed for traversal. Taking traversal in range of any two parameters, the model fixes the rest seven parameters and analyze three corresponding eigenvalues of three sets of real solutions.

Select k_1' traversing in [0.01, 1] and select k_3 traversing in [0.01, 20] (Figure 5.4). As seen from Figure 5.4, three corresponding eigenvalues of three sets of real solutions are separately shown in A, B, C. The horizontal axis represents the value of k_1' and the vertical axis represents the value of k_3. Similar with the single

parameter traversal discussion, there also exists situations which the number of stable points will be 0, 1 and 2. The two parameters traversal analysis of other parameters are similar with the above discussion in Figure 5.4. The detailed data can be found in the appendix III.

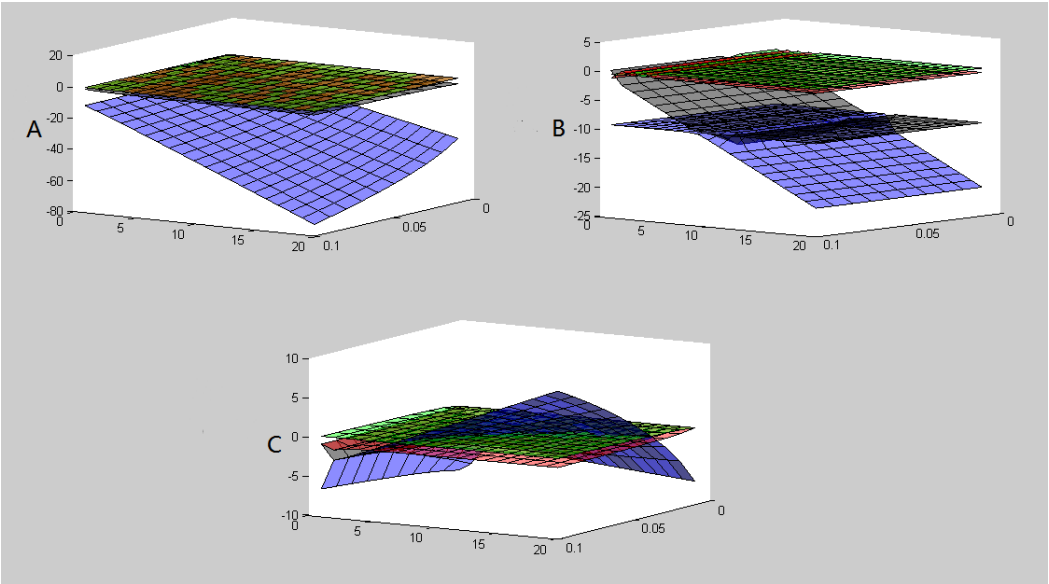


Figure 5.4 Three corresponding eigenvalues of three sets of real solutions are separately shown in A, B, C when k_1' take traversal in [0.01, 1] and k_3 traversing in [0.01, 20]. (The green plane is the concentration of the INSR, the red plane is the concentration of the LAR, the blue plane is the concentration of the GRB2, the black plane is the concentration of the IRS)

By analyzing the resulting data, some views can be shown as follow:

1) The parameters range with two stable steady points is approximately continuous. But in some specific interval, the numbers of stable steady points are interval oscillation.
2) The situation with three stable steady points is not obtained in the model. Maybe there does not exist three stable steady points in the system.

3) The percentage of the situations with two stable steady points is higher than 70%. So there is a high possibility that can obtain more than one stable steady point.

5.5 Conclusion

Type 2 diabetes is a long-term metabolic disorder that is characterized by high blood sugar, insulin resistance, and relative lack of insulin. Rates of type 2 diabetes have increased markedly since 1960 in parallel with obesity [Mos10]. As of 2015 there were approximately 392 million people diagnosed with the disease compared to around 30 million in 1985 [Smy06]. Typically it begins in middle age [Gao14] although rates of type 2 diabetes are increasing in young people ([Tfa09], [Imp12]). Type 2 diabetes is associated with a ten-year-shorter life expectancy. There are two possible reasons for insulin resistance. Firstly, the concentration of insulin may be different and the external stimulation is also different. The activator protein is to add a phosphoric acid group on protein, which is also called phosphorylation. With the different regulations of enzyme, the phosphorylation levels are different. Secondly, there may be many receptors on the surface. With the different numbers of active receptors and the different stimulations, activation statuses are different. It is hard to make such kind experiment under the single cell level. The pertinent reaction of the single cell to the insulin is hard to obtain.

There are many causes of diabetes, such as obesity, lack of exercises or genetic resistance of insulin. The therapies of type II diabetes mainly depend on insulin

injections, but there always exists the resistance to insulin in the process of the treatment. This research built the model based on the proteins in the insulin signaling pathway. Through the simulation and computing, the model may have more than one stable steady point under the same external stimuli. So at the same concentration of insulin stimulation, the cell can reach the stable steady state under two different conditions. As for the proteins in the model, GRB2 can regulate and control downstream protein synthesis or cell proliferation through the MAPK signaling pathway. IRS can regulate and control blood sugar by the regulation of GLUT4. Moreover, the activations of some proteins can make an impact on lipid metabolism in the process of active glycogen synthesis. Consequently, two stable steady states can represent two different states of cell itself and peripheral extracellular matrix, also two different states of organisms. This research proposed the discussion of multiple homeostasis problems in the dynamic process of insulin receptor pathway and considered that the two different states may be the significant causes of insulin resistance. Previous studies have shown that more than one steady states exist in cell cycle model and DNA replication model. Shift between different steady states happens during biological process. Although there is no evidence show that two states can be free to switch or transform each other under certain external stimuli, the possibility proposed in this study can provide reference for further researches of insulin signaling pathway.

Chapter 6 Meta-Analysis of Gene Expression during Heart Regeneration in *Danio rerio*

6.1 Introduction

In recent years, regenerative medicine has gained increasing attention. As a model vertebrate, *Danio rerio* (zebrafish) has appreciable regeneration ability for the retina, fish fin, heart, and other tissues and organs, thereby becoming an important model for studying organ regeneration. The recent discovery of cardiac regeneration ability and preliminary discovery of the underlying molecular and cellular mechanisms have been major breakthroughs in the field of organ regeneration. Although mammalian hearts were widely believed to be unable to regenerate in 2002, Poss et al. [Pos02] reported, for the first time, that after surgical removal of approximately 20% of the ventricle of an adult *Danio rerio*, the wound was almost completely repaired within two months. Inspired by the report of *Danio rerio* in 2011, Porrello et al. [Por11] reported that following birth, a mouse heart could regenerate after seven days of operation but lost this ability after seven days. Therefore, studying the molecular mechanism of adult heart regeneration in *Danio rerio* would help restore the ability of adult heart regeneration in mammals, and provide a theoretical basis for the treatment of heart diseases such as myocardial infarction. Currently, heart regeneration in *Danio rerio* has become a hot spot in the field of organ regeneration research in adult vertebrates; researchers are participating in the regeneration of cell lineage and adult cardiomyocytes, and a

series of valuable results have been obtained regarding their underlying molecular mechanisms ([Ber09], [Zhe12]). Three methods are reported for inducing heart regeneration in *Danio rerio* by artificial injury: surgical resection, freezing injury, and toxin treatment. Freezing injury is generated by frostbite and ventricular necrosis, caused by direct contact of the heart with a metal probe treated with liquid nitrogen. The resulting damaged area is difficult to control, and the cycle of heart regeneration is relatively long [Gon11]. It is the cardiomyocyte-specific promoter that controls the expression of diphtheria toxin DTA by conditioning the CreER/loxP system, in parts of the adult heart, to achieve the heart injury [Wan11]; however, this method does not cause acute injury. The process of cell necrosis and scar formation may be different for surgical excision or freezing injury in regeneration. Compared to the above two methods, surgical resection for cardiac injury is easier to operate and handle.

The feasibility of genetic manipulation and corresponding techniques has enabled the study of vertebrate organ regeneration. Besides their myocardium, fins, retina, optic nerve, liver, spinal cord, and sensory hair cells, *Danio rerio* demonstrated the ability to reconstruct damaged heart tissue after partial cardiac excision. Poss et al. [Por11] screened cell cycle-regulated mitotic checkpoint protein, (monopolar spindle 1, MPS1), which is necessary for heart and fin regeneration in *Danio rerio*. In addition, regeneration involves multiple signaling pathways, including family members of fibroblast growth factor, FGF, and Notch signaling pathways ([Ray03], [Lep06]), growth factors, secretory factors, and others involved in the regulation

of inflammation, cell adhesion or extracellular matrix factors. Lien et al. [Lie06] and Sleep et al. [Sle10] reported noticeable changes in heart regeneration and the surrounding region in *Danio rerio*. Many similarities exist between the heart mutants in *Danio rerio* and heart diseases in humans ([Ger02], [XuX02]). For example, mutations in *Danio rerio* and human TTN gene, which encodes for the giant muscle filament titin, can cause cardiomyopathy ([LiQ97], [Gar02]). Mutation in hERG (human Ether- à -go-go related gene), also known as K^+ gated channel H subclass 2, KCNH2, may lead to arrhythmia ([Cur95], [Lan03]). Based on these findings, it was assumed that the heart regeneration process of *Danio rerio* can be used as a model to study the mechanism of human heart diseases and develop a treatment to help overcome heart diseases.

Heart regeneration in *Danio rerio* is a highly dynamic process. The events that occur soon after the injury have an important impact on myocardial regeneration and interaction between cardiomyocytes and non-cardiomyocytes. For the sake of simplicity, the regeneration process was subdivided into the following stages: 1) early response to injury (including inflammation and endocardial activation; 2) endocardial and epicardial regeneration; 3) cardiomyocyte proliferation; and 4) integration of regenerated cardiomyocytes into the myocardium and removal of scars. The study illustrated the process in response to cryoinjury as an example, although the information from different damage models have been collected.

At different stages of cryopreservation, local tissue necrosis (gray) and apoptosis of all cell types around the injured area are induced. Tissue death triggers the

recruitment of inflammatory cells and activation of the endocardium. In the first few days after injury, epicardial and endocardial cells proliferate actively and cover the damaged area to establish "regenerative stents." Epicardial cells also undergo epithelial-to-mesenchymal transitions and invade the underlying myocardium. Myofibroblasts appear in damaged areas and have extracellular cells that accumulate as the matrix, proliferate at the edge of the wound, and rebuild the damaged area. With myocardial regeneration, fibrotic tissue gradually disappears. At the later stage of regeneration, *Danio rerio* myocardium shows complete recovery. The gene sequence of zebrafish is about 70% same with the human, so this study proposed to find the gene expression of zebrafish to analyze the human function better. Compared to uninjured control or contralateral wall, the regenerated wall shows significant dilatation of cortical myocardium; however, molecular mechanism underlying heart regeneration is not clear enough. In this study, the gene expression profiles related to cardiac regeneration was obtained through meta-analysis and analyzed to obtain differentially expressed genes, which provided a reference for future research in related fields.

6.2 Material and Method

Biological information

Sequencing data from all published articles that contain RNA-seq data relevant to zebrafish heart regeneration are collected ([Han14], [Kan16], [Rod16], [WuC16]). GRCz10 Ensembl release 91 was used as reference genome and analyzed process

described below. Sequencing results were downloaded through sra-toolkit (https://www.ncbi.nlm.nih.gov/sra/docs/toolkitsoft/), hisat2 software (https://ccb.jhu.edu/software/hisat2) was used to align the sequence on the genome, StringTie software (https://ccb.jhu.edu/software/stringtie/) was used to splice, and finally R package ballgown (https://www.bioconductor.org/packages/3.7/bioc/html/ballgown.html) was used to summarize and visualize the results.

Clustering via nonparametric distribution estimation:

Clustering method

Clustering analysis is the most commonly used method to deal with unlabeled data, the study attempted to cluster the data set to find out whether there was a distinct subset of data. First, the data whose fold change (FC) value was 1 are excluded, since both the p-value and q-value are infinite, and then the correlation coefficient matrix of 3D data was calculated:

$$C = \begin{pmatrix} 1.0000 & -0.2231 & -0.2482 \\ -0.2231 & 1.0000 & 0.9452 \\ -0.2482 & 0.9452 & 1.0000 \end{pmatrix} \quad (6.1)$$

K-means and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), since the second and third dimensions of the data had a very strong positive correlation, in the following analysis, the study abandoned the third dimension of the result and used only the first two dimensions for analysis. The mainstream clustering methods are divided into two main categories: distance-based clustering and density-based clustering. By comparing the distance between

the data points and the center of each class (usually Euclidean distance), the distance-based clustering determines which cluster these data points belong to. The density-based clustering is based on the density of data points (such as the number of data points in a given range). In this work, two classical clustering methods were used: K-means and DBSCAN.

The principle of K-means clustering is based on the known observational set $X = \{x_1, x_2, x_3, \dots, x_n\}$. It is divided into k sets, so that the sum of squares in the group is minimum; that is, the partition $S = \{S_1, S_2, S_3, \dots, S_k\}$ so that

$$S = \underset{S}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2 \tag{6.2}$$

where μ_i is the mean value of all the points in the class of S_i.

DBSCAN (Density-based spatial clustering of applications with noise) is a density-based clustering method. Contrary to the K-means algorithm, which calculates the distance between the sample points and k-class centers, this algorithm divides the sample points close enough into a class and constructs number-density regions based on the sample points divided into classes. Its greatest advantage over K-means is the ability to learn non-spherical data sets.

Nonparametric distribution estimation

The distribution estimation method usually used always presupposes that the data conform to a given distribution and subsequently estimates the parameters of distribution using statistical analysis of the data. However, the current analysis in this study shows that the data do not fit the common distribution model. Hence,

the nonparametric method is used to estimate the joint probability density function. Let PX be the probability density function of a two-dimensional random variable; the probability of random sampling in a region of sample space R is:

$$P_R = \int_R p(x,y)dxdy \tag{6.3}$$

When the sample size N is large enough, and the volume of R is V, it can approximately be considered as

$$P_R = \mathrm{p(x,y)V} \tag{6.4}$$

In this case, if k samples fall in the region R, the probability density at the central point of region R is estimated as:

$$\hat{P}(x_0, y_0) = \frac{k/N}{V} \tag{6.5}$$

The objective of this estimation is to count the data points in the region with volume V near point (x_0, y_0). Obviously, when the sample size N is determined, selection of volume V has a great effect on the estimation. Therefore, it is assumed that shrinking regions R1, R2, R3, and R4, are constructed when N is infinite. RN for continuous estimation of P(x_0, y_0)

$$\hat{P}_n(x_0, y_0) \approx \frac{k_n/N}{V_n} \tag{6.6}$$

When the following three conditions are met, $\hat{\mathrm{P}}_\mathrm{n}(\mathrm{x_0,y_0})$ will converge on $\mathrm{P(x_0,y_0)}$.

1. $\lim_{n\to\infty} V_n = 0$: the region will eventually converge to the point (x_0, y_0)

2. $\lim_{n\to\infty} k_n = \infty$: there are always enough points to estimate in the region

3. $\lim_{n\to\infty} k_n/N = 0$: the number of points in the region is only a fraction of the overall data.

The Parzen window method is most frequently used to calculate the k_n included in the window by the given relationship between V_n and n. The commonly used window functions are rectangular window, Gaussian window, exponential window, etc.

The volume of the selected region as per estimate:

$$V_n = h/\sqrt{n} \tag{6.7}$$

The Gaussian function satisfies the sampling center with a higher weight than the far position and is continuously steerable throughout the window region with good mathematical properties. Considering the ease of use of functions in waveform synthesis calculations, Gaussian functions (normal distribution density functions) are generally used as window functions:

$$\varphi(\mathrm{x},\mathrm{y}) = \frac{1}{2\pi}\exp\{-\frac{(x^2+y^2)}{2}\} \tag{6.8}$$

Therefore, the final estimate of probability density

$$\hat{P}(x,y) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{V_n}\varphi(\frac{x-x_i,y-y_i}{V_n}) \tag{6.9}$$

6.3 Results

By analyzing the sequencing results from previous literature, the expression of

more than 30,000 genes in *Danio rerio* was obtained during cardiac regeneration. The fold change data on the left side of the heat map above indicated that left column of the heatmap indicates fold change data, where green lines represent genes upregulated in the regeneration group; on the contrary, red lines represent genes downregulated in the regeneration group (Figure. 6.1). Overall, the number of genes up-regulated was similar to the down-regulated ones. Furthermore, functional cluster analysis for up-regulated and down-regulated genes showed that the up-regulated genes were mainly enriched in oxidative phosphorylation, metabolism, amino acid degradation, glycogen synthesis, protein nucleation, spindle formation, and other biological processes, while down-regulated genes were mainly concentrated in the p53 signaling pathway, lysosome, protein degradation, and other processes. During the process of heart regeneration in *Danio rerio*, numerous energy-related genes are up-regulated, indicating that regeneration requires a large amount of energy, probably due to the increased requirement of protein synthesis and cell division. Up-regulation of protein entry indicates that transcription factors enter the nucleus and regulate the transcription and expression of downstream-related genes. The most direct evidence indicating the entry of cells into the cell cycle is the formation of spindle bodies. Down-regulation of gene clustering is mainly found in apoptosis. Protein degradation and other processes were also found as per expectations. Figure 6.1 shows changes in gene expression during heart regeneration in *Danio rerio*; green indicates up-regulated, and red indicates down-regulated. Figure 6.2 shows gene ontology (GO)

cluster analysis of up-regulated genes during cardiac regeneration. Figure 6.3 shows GO cluster analysis of down-regulated genes during cardiac regeneration. Digging into clustering results, oxidative phosphorylation is significantly increased, which consumes massive ATPs produced by glycolysis and other relative reactions. Up-regulated metabolic pathways consistent with this phenomenon. However, no metabolism-related pathways clustered in down-regulated genes. All in all, carbon metabolism could play an important role during zebrafish heart regeneration. Figure 6.4 shows distribution of differentially expressed genes during heart regeneration in *Danio rerio,* the gene represented by the red dot is the gene with the most significant difference in expression. Seven significantly changed genes are found. Among them, mrc1b and aox6 are relevant to carbon metabolism, while pvalb7 and atxn2l are crucial for cell signal delivery.

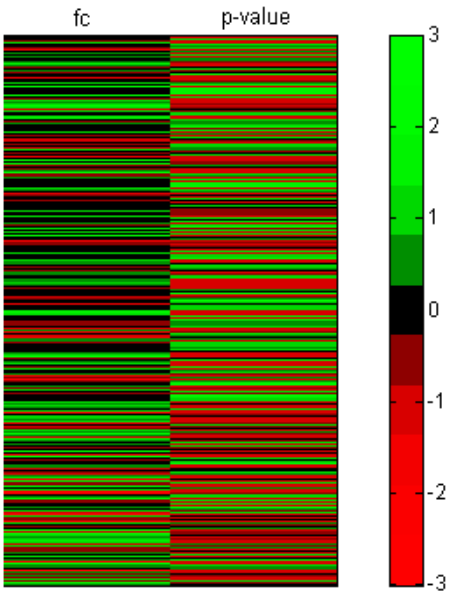


Figure 6.1 Changes in gene expression during heart regeneration in *Danio rerio*;

green indicates up-regulated, and red indicates down-regulated.

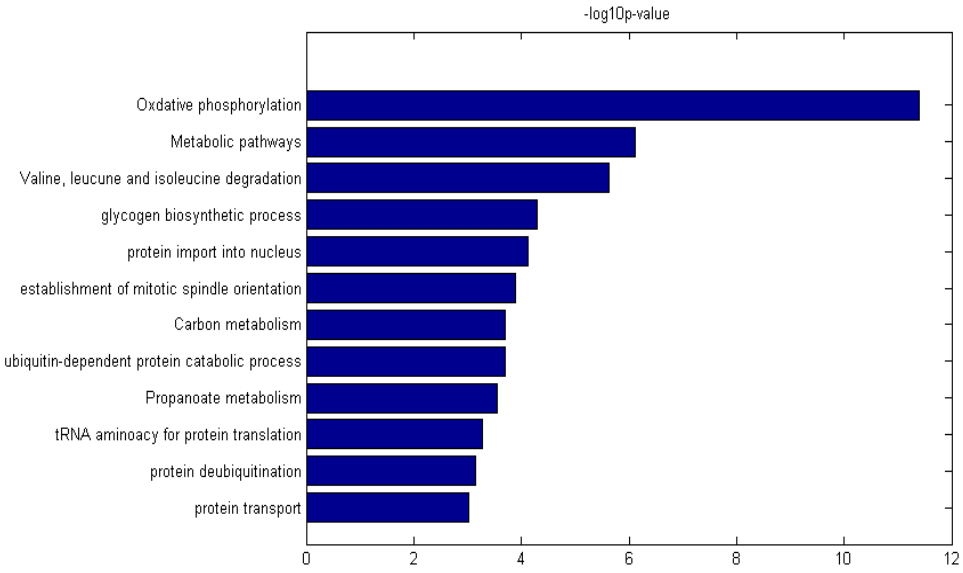


Figure 6.2 GO cluster analysis of up-regulated genes during cardiac regeneration;

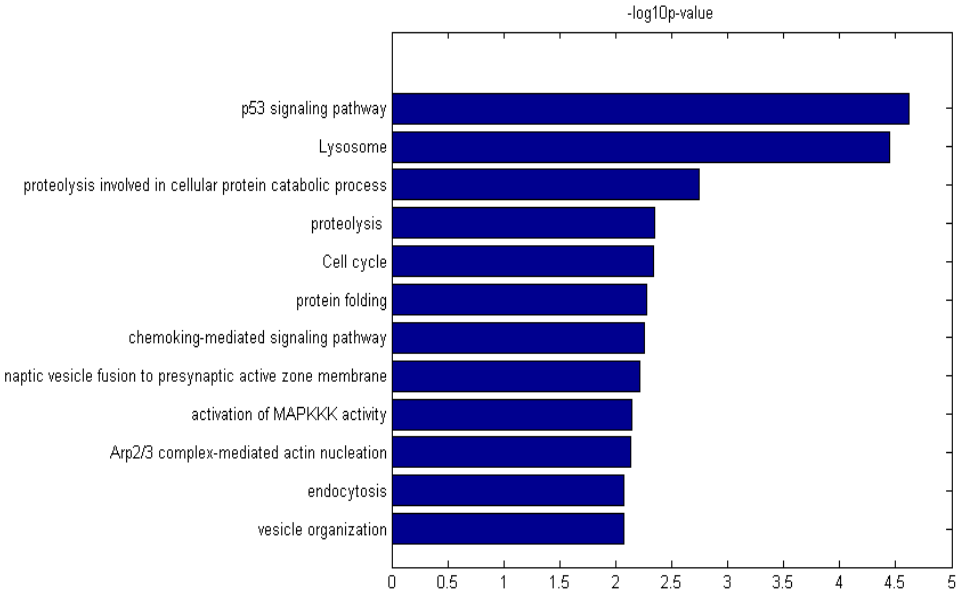


Figure 6.3 GO cluster analysis of down-regulated genes during cardiac regeneration;

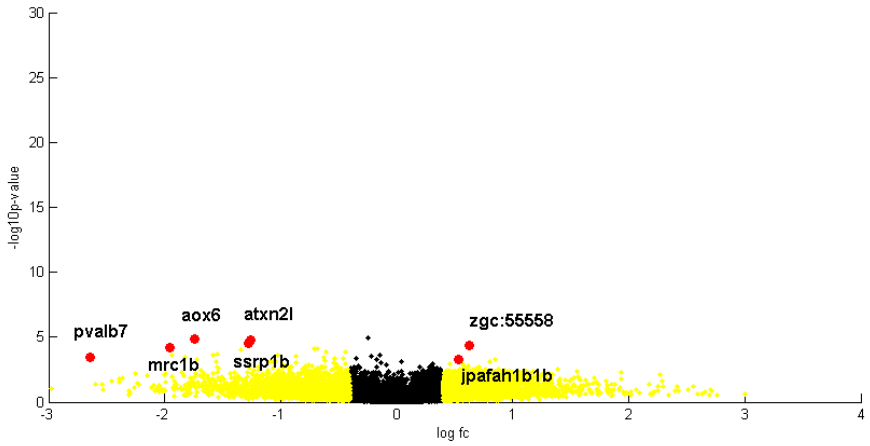


Figure 6.4 Distribution of differentially expressed genes during heart regeneration in *Danio rerio,* the gene represented by the red dot is the gene with the most significant difference in expression.

Cluster analysis

The K-means algorithm has no mechanism to specify the number of categories. After comparing the distance between the final clustering centers, it is best to divide the data into six categories. However, since the K-means method was not suitable for use, DBSCAN was used. It is a classical density-based clustering method. The result is shown below; almost all the data were classified into one class by the algorithm. However, the data points that did not belong to this class, including the outliers showed the whole data to be very polymeric, and it could not be divided by clustering analysis (Figure.6.5).

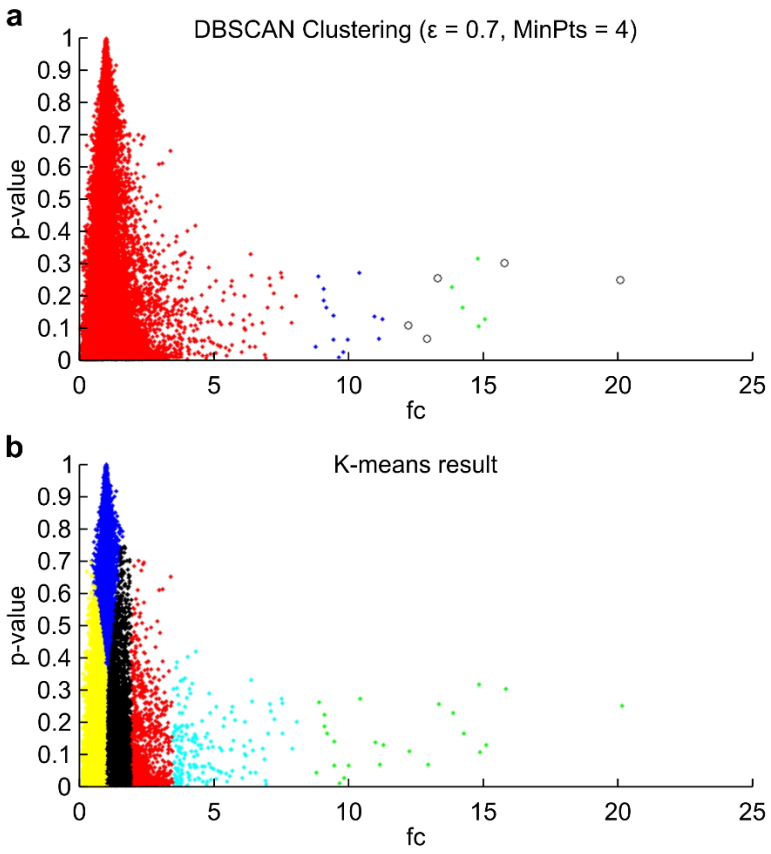


Figure 6.5 (A): DBSCAN clustering results of FC-p-value; (B): K mean clustering results of FC-p-value.

Distribution estimation:

Statistical method was used to estimate the distribution of the data. Firstly, calculate the digital characteristics of the two-dimensional data.

Expected:

$$E_{fc} = 1.1430, E_{pvalue} = 0.3888 \tag{6.10}$$

Covariance matrix:

$$\text{Cov} = \begin{pmatrix} 0.5405 & -0.0477 \\ -0.0477 & 0.0845 \end{pmatrix} \tag{6.11}$$

A histogram of two edge distributions is shown.

The marginal distribution showed that the distribution of neither of FC or p-values satisfied the Gaussian distribution; therefore, the joint distribution difficult to be estimated by the classical multivariate Gaussian distribution.

Parzen window estimation

The joint probability density of the data set is estimated to be

$$\hat{P}(x,y) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{V_n}\varphi\left(\frac{x-x_i, y-y_i}{V_n}\right) \quad (6.12)$$

Where,

$$\varphi(\mathrm{x},\mathrm{y}) = \frac{1}{2\pi}\exp\left\{-\frac{(x^2+y^2)}{2}\right\} \quad (6.13)$$

and,

$$V_n = h/\sqrt{n} \quad (6.14)$$

is the volume of selected region in each estimation. According to the data collected, the value of *h* is 20 (Figure. 6.5).

Figure 3 near here

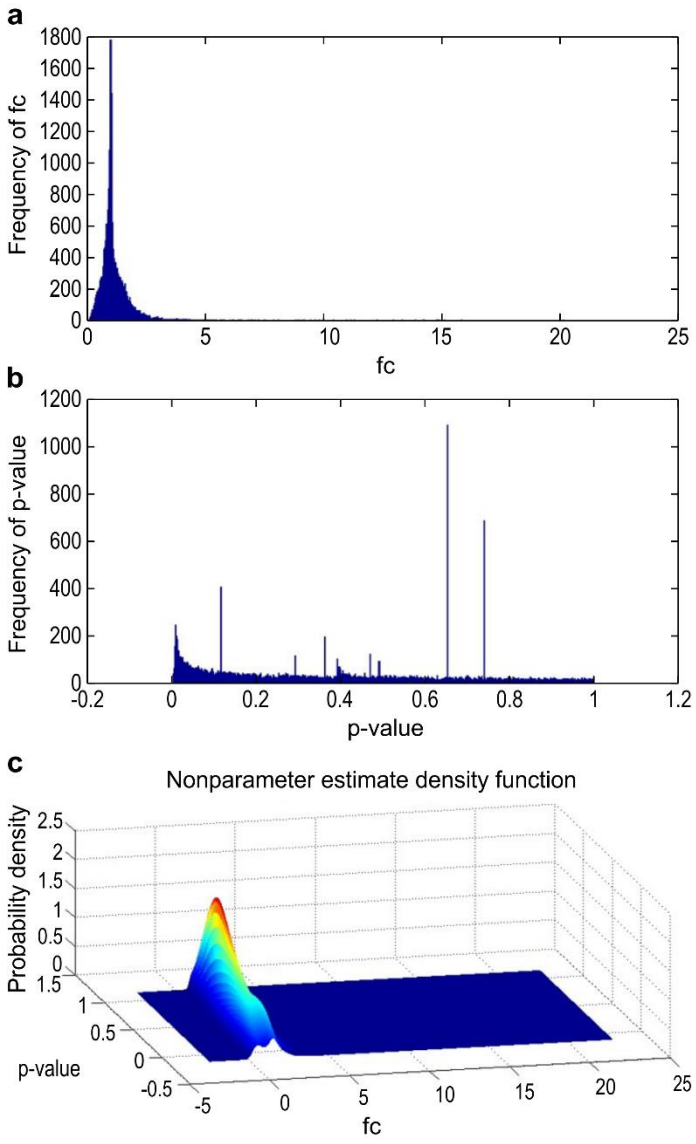


Figure 6.6 (A): FC histogram; (B): p-value histogram; (C): FC p-value joint distribution probability density.

6.4 Discussion

In this study, all the published RNA-seq results of heart regeneration in *Danio rerio* is collected and analyzed again. The expression data of more than 30,000 genes in *Danio rerio* during heart regeneration is obtained. The results were analyzed by genetic clustering and mathematical statistics. Genetic cluster analysis showed that

heart regeneration involves a large number of energy-intensive biological processes that may include new protein synthesis, transcription of new genes, cell division, and immune response, while spindle formation indicated that cells enter the cell cycle process during the regeneration, thereby validating the observation made earlier by Poss et al. [Pos02].

In the corresponding down-regulated gene cluster, apoptosis and protein degradation were inhibited, suggesting that the regeneration process does not rely solely on cell division to produce new cells; rather, the old cell apoptosis rate may slow down, further accelerating the regeneration process. It was also noted that the down-regulated cluster has some cell cycle-related genes, possibly due to a specific cell proliferation mechanism which needs to be suppressed during regeneration. Based on the observation, it would be convenient for the subsequent researchers to screen and analyze the results further.

Through meta-analysis, using more data, more accurate gene differential expression profiles during zebrafish heart regeneration and the corresponding biological processes by go-cluster analysis are obtained. Compared with the analysis results of any of them, our analysis is more comprehensive and reliable and excludes experimental deviations caused by human operations.

Chapter 7 Conclusions

7.1 Overall Conclusions

In the post-genomic era, the resolution of genomic functions has become a hotspot in life science research. Based on the complexity of gene function and the integrity of biological systems, it is necessary to understand the functions of the various modules that make up the biological system on a holistic level. With extensive biological research and accumulation of biological information, coupled with the development of new measurement techniques, high-throughput analytical methods, advanced information science and new theories of systems science, it is possible to study the life system formed by the discovery of molecular biology at the system level. Life sciences has once again reached at the new height of integrated research, gradually entering the era of systems biology from the era of molecular biology. Therefore, genomics-derived transcriptomics, proteomics, and metabolomics as holistic research methods have become new and important tools for the analysis of functional genomics. Metabolomics studies the quality, quantity and the changes in various metabolites produced by biological systems after intervention by external factors. Since the metabolite is at the end of the regulation of biochemical activities in the biological system, it contains direct and comprehensive biomarker information reflecting the physiological phenotype. Metabolomics is increasingly becoming a powerful analytical tool for studying the functional changes of life systems in a holistic manner.

A biological organism is a complex system with dynamic, multi-factor integrated regulation. In the biological information transmission chain from genes to traits, the body needs to constantly adjust its complex metabolic network to maintain the normal dynamic balance within the system and the external environment. The DNA, mRNA, and proteins provide a material basis for the occurrence of biological processes, while metabolites and metabolic phenotypes reflect the biological events that have occurred, the combined results of genotypes and the environment and direct expression of physiological and biochemical functional status of biological systems. Therefore, as an important component of systems biology, the metabolome can better reflect the system phenotype. The function of cells is largely reflected in the regulation on metabolic levels. The expression profiles of some related genes are similar, but the metabolite profiles or metabolic fluxes are significantly different. Fluxome is a terminal embodiment of genomic function. In case of a disturbance, the final metabolic flux remains relatively stable to counteract the interference effect due to the compensation of genes or proteins. Metabolic flux is a fundamental determinant of cell physiology and the most important parameter in metabolic pathways. Metabolic flux analysis (MFA) is a method for determining the metabolic flux distribution in the entire metabolic reaction network based on the measurement of the relationship between each reaction and the metabolic pathway and the data measured in the experiment and plays an important role in metabolic engineering. By calculating the metabolic flux distribution under different pathways or different conditions, the metabolic

capacity of the cells can be characterized, and the influence of genetic modification on the metabolic state of the cells can be observed, thus providing a theoretical basis for further rational genetic transformation. MFA combines data on substrate uptake and product formation rates, biosynthesis requirements, stoichiometric reactions, and biosynthesis of intermediate metabolites to provide critical quantitative information for reconstituting cellular metabolism. The pathways that MFA focuses on are usually central carbon metabolic pathways (including glycolytic pathways, tricarboxylic acid cycle, and pentose phosphate pathway), and sometimes include partial biosynthetic pathways (such as amino acid biosynthesis). MFA has been successfully applied to the optimization of various biological metabolic pathways, and the yield of various target products including organic acid vitamins and ethanol has been improved. In this research, an innovative and generic approach has been proposed to analyze metabolic flux. All the data are collected from published data. The approach combined metabolic flux analysis, continuous-time Markov chain and Monte Carlo method and is applied to central metabolic pathways. This study also applied the MFA approach to other disease models related to metabolism. The insulin signaling pathway was selected as the process model to analyze the biological characteristics in insulin resistance when the system reaches steady state. It is found that the multiple homeostasis problems in the dynamic process of insulin receptor pathway and considered that the two different states may be the significant causes of insulin resistance. Finally, meta-analysis is applied to the gene expression during heart regeneration in Danio

rerio to find several genes that have not been linked to the heart regeneration process that might be useful for rebuilding cardiac function in zebrafish. The approaches in this study can also be applied on other diseases models.

7.2 Contributions

The contributions of this research are as follows:

(1) An innovative and generic approach, which combines metabolic flux analysis and Continuous-time Markov Chain, has been put forward to analyze metabolic flux. Based on the study of the pentose phosphate pathway, this approach calculated the steady-state concentration by the distribution of each carbon atom. The research obtained the reaction equilibrium constant for analysis and comparison, and found this approach is feasible and the results are accurate. Furthermore, the combination of the Continuous-time Markov chain method and Monte Carlo method was applied to the central pathways in glucose metabolism. This study adopted the Monte Carlo method to solve the problem that the reaction rate was constant is not available and to calculate the equilibrium concentration with the distribution of each carbon atom when all the reactions reach steady state. The results show that the method is feasible; and the results are accurate. The approach is generic and can be applied to any metabolic system.

(2) Application of the MFA approach to the insulin signaling pathway, revealed multiple homeostasis problems in the dynamic process of the insulin receptor pathway. Diabetes has many causes, such as obesity, lack of exercise, or genetic

resistance to insulin. The treatment of type 2 diabetes mainly relies on insulin injection, but resistance to insulin always exists during the treatment. The proposed model is based on proteins in the insulin signaling pathway. Through simulation and calculation, the model can have multiple stable points under the same external stimulus. Thus, at a same insulin stimulation concentration, cells can achieve stable homeostasis under two different conditions. Considering the proteins in the model, GRB2 can regulate and control downstream protein synthesis or cell proliferation through the MAPK signaling pathway. The IRS proteins can regulate and control blood sugar by regulating GLUT4. In addition, the activation of some proteins may affect lipid metabolism in the process of active glycogen synthesis. Thus, two stable homeostasis points can represent two different states of the cell itself and the extracellular matrix, as well as two different biological states. This study presents a discussion of multiple homeostasis problems in the dynamic process of insulin receptor pathways and suggests that these two different states may be important causes of insulin resistance. Although there is no evidence that the two states can switch freely or interchangeably under certain external stimuli, the possibility presented in this research can provide reference for further study of the insulin signaling pathway.

(3) Meta-analysis was applied to gene expression during heart regeneration in *Danio rerio*. All published RNA-seq results from *Danio rerio* cardiac regeneration are collected and reanalyzed. The expression data of more than 30,000 genes in *Danio rerio* are obtained during cardiac regeneration. The results were analyzed

by genetic clustering and mathematical statistics. Genetic cluster analysis showed that cardiac regeneration involves a large number of energy-intensive biological processes, possibly including new protein synthesis, new gene transcription, cell division and immune response, while spindle formation indicates that cells enter the cell cycle during regeneration. In the down-regulated gene clusters, apoptosis and protein degradation were inhibited, indicating that the regeneration process was not only dependent on cell division to produce new cells. In contrast, the rate of old cell apoptosis may slow down, further accelerating the regeneration process. It was also noted that the downregulated clusters had some cell cycle related genes, possibly due to specific cell proliferation mechanisms that needed to be suppressed during regeneration. Based on our observations, follow-up researchers can easily further screen and analyze the results. Through meta-analysis and using more data, more accurate gene differential expression profiles in the regeneration process of the zebrafish heart and corresponding biological processes were obtained through cluster analysis. Our analysis is more comprehensive and reliable than any of the results, and excludes experimental biases caused by human manipulation.

7.3 Limitations of the research

In this dissertation, a dynamic MFA method is proposed and this study attemptes to apply it to complex processes of metabolic models and metabolic disease-related models. However, there are still some limitations need to be overcome.

(1) All the reaction mechanisms and parameters of the biological system process

need to be available, but some of them, such as reaction rate parameters, difficult to be completely obtained, so they can only be replaced by the Monte Carlo sampling method.

(2) Compounds in each system process are also associated with other system processes, which can lead to errors in the analysis of individual process models. For example, the corresponding metabolite F6P is also part of fructose and mannose metabolism, which generates GDP in the central metabolism process for further glycosylation.

(3) Some parameters and control data can be only obtained through the literature, and the data is difficult to be controlled without deviation.

7.4 Future work

It is believed that the research conducted in this dissertation provides researcher an opportunity to continue contributing towards the development of metabolic flux analysis. Following suggestions are proposed for future work:

(1) The continuous-time Markov chain MFA method should be applied to a larger and more complex biological system model in order to expand the application scope of this method and enhance the effectiveness of the method and the accuracy of the results.

(2) By adjusting the method, the model can be applied to a multi-stable point biological system process, which can effectively analyze the disease model which is similar to insulin resistance and provide more useful information for the analysis of diabetes.

(3) Combining the above model with the metabolic data of cardiovascular patients, the causes of cardiovascular diseases and drug resistance can be analyzed from the perspectives of glucose metabolism and insulin, which can be further extended to other metabolic disease models such as cancer.

Appendices

Appendix I Example of the Continuous-time Markov Chain approach

Calculate our model in detail as follows. The key point is to calculate the Q-matrix, combining with the equations (4) (5) (6) and (7) in section 2, then obtain:

$$q_{ab}^{i}(\infty)=\begin{cases}\dfrac{v_{a\to b}^{i}(\infty)}{[A_a(\infty)]} & a\neq b\\ -\sum\limits_{b=1,b\neq a}^{m} q_{ab}^{i}(\infty) & a=b\end{cases}$$

For example, the reactions containing the transfer processes (from A_1 to A_2) of the 1st carbon atom existed in A_1 is:

$$A_1 \overset{k_1}{\to} A_2$$

So

$$v_{1\to 2}^{1}(\infty)=k_1[A_1(\infty)]$$

$$q_{12}^{1}(\infty)=\frac{v_{1\to 2}^{1}(\infty)}{[A_1(\infty)]}=k_1$$

As the same way, the 1st line of $Q^1(\infty)$ is:

$$q_{12}^{1}(\infty)=k_1, q_{13}^{1}(\infty)=k_3, q_{14}^{1}(\infty)=\cdots= q_{18}^{1}(\infty)=0,$$

$$q_{11}^{1}(\infty)=-\sum_{b=2}^{8} q_{1b}^{1}(\infty)=-k_1-k_3$$

So,

$q_{21}^{1}(\infty)=k_2,$

$q_{22}^{1}(\infty)=-k_2-k_5[A_3(\infty)]u,$

$$q_{24}^{1}(\infty) = q_{25}^{1}(\infty) = 0.5k_5[A_3(\infty)]u,$$

$$q_{31}^{1}(\infty) = k_4,$$

$$q_{33}^{1}(\infty) = -k_4 - k_5[A_2(\infty)]u - k_9[A_7(\infty)]u,$$

$$q_{34}^{1}(\infty) = 0.5k_5[A_2(\infty)]u,$$

$$q_{35}^{1}(\infty) = 0.5k_5[A_2(\infty)]u + 0.5k_9[A_7(\infty)]u,$$

$$q_{36}^{1}(\infty) = 0.5k_9[A_7(\infty)]u,$$

$$q_{42}^{1}(\infty) = q_{43}^{1}(\infty) = 0.5k_6[A_5(\infty)]u,$$

$$q_{44}^{1}(\infty) = -k_7 - k_6[A_5(\infty)]u,$$

$$q_{47}^{1}(\infty) = q_{48}^{1}(\infty) = 0.5k_7,$$

$$q_{52}^{1}(\infty) = 0.5k_6[A_4(\infty)]u,$$

$$q_{53}^{1}(\infty) = 0.5k_6[A_4(\infty)]u + 0.5k_{10}[A_6(\infty)]u,$$

$$q_{55}^{1}(\infty) = -k_6[A_4(\infty)]u - k_{10}[A_6(\infty)]u - k_{12}[A_8(\infty)]u - k_{14},$$

$$q_{56}^{1}(\infty) = k_{12}[A_8(\infty)]u,$$

$$q_{57}^{1}(\infty) = 0.5k_{10}[A_6(\infty)]u,$$

$$q_{58}^{1}(\infty) = k_{14},$$

$$q_{63}^{1}(\infty) = q_{67}^{1}(\infty) = 0.5k_{10}[A_5(\infty)]u,$$

$$q_{65}^{1}(\infty) = q_{68}^{1}(\infty) = 0.5k_{11},$$

$$q_{66}^{1}(\infty) = -k_{10}[A_5(\infty)]u - k_{11},$$

$q_{74}^{1}(\infty) = k_8[A_8(\infty)]u,$

$q_{75}^{1}(\infty) = q_{76}^{1}(\infty) = 0.5k_9[A_3(\infty)]u,$

$q_{77}^{1}(\infty) = -k_8[A_8(\infty)]u - k_9[A_3(\infty)]u,$

$q_{84}^{1}(\infty) = k_8[A_7(\infty)]u,$

$q_{85}^{1}(\infty) = k_{13},$

$q_{86}^{1}(\infty) = k_{12}[A_5(\infty)]u,$

$q_{88}^{1}(\infty) = -k_8[A_7(\infty)]u - k_{13} - k_{12}[A_5(\infty)]u,$

The rest entries of the matrix are 0.

Mentioned in the reseach, $Q^1 = Q^2 = Q^3 = Q^4 = Q^5$, so $\pi_j^1 = \pi_j^2 = \pi_j^3 = \pi_j^4 = \pi_j^5$ (j=1, 2,…, 8)

By equation (11), calculate $[A_j(\infty)]$

For example,

$$[A_1(\infty)] = \frac{(\pi_1^1 + \pi_1^2 + \cdots + \pi_1^5)}{N_1} = \frac{5\pi_1^1}{N_1} = \frac{5\pi_1^1}{5} = \pi_1^1$$

Similarly,

$$[A_2(\infty)] = \pi_2^1,$$

$$[A_3(\infty)] = \pi_3^1,$$

$$[A_4(\infty)] = \frac{5}{7}\pi_4^1,$$

$$[A_5(\infty)] = \frac{5}{3}\pi_5^1,$$

$$[A_6(\infty)] = \frac{5}{6}\pi_6^1,$$

$$[A_7(\infty)] = \frac{5}{4}\pi_7^1,$$

$$[A_8(\infty)] = \frac{5}{3}\pi_8^1,$$

Put above 8 equations into the matrix Q, and combine the matrix with equation (8), π_j^h can be obtained by the MATLAB. Putting π_j^h into equation (11), the relative concentrate $[A_j(\infty)]$ can be easily obtained. Then K can be calculated with the relative concentrates.

For example, assume *u*=0.02,

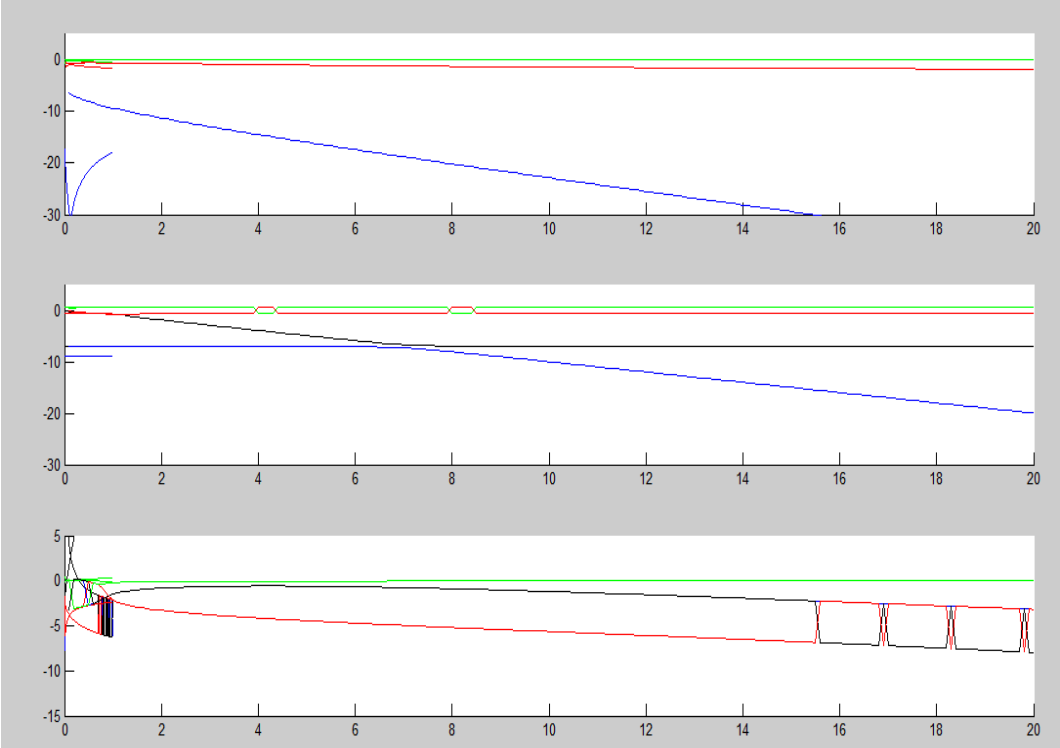
$(\pi_1^1, \pi_2^1, \ldots, \pi_8^1)$=(0.078799, 0.033975, 0.110265, 0.297409, 0.014218, 0.170622, 0.001896, 0.292816) (there is only one set of positive solutions)

$$K_1 = \frac{[A_3(\infty)]}{[A_1(\infty)]} = \frac{\pi_3^1}{\pi_1^1} = \frac{0.110265}{0.078799} = 1.3993198$$

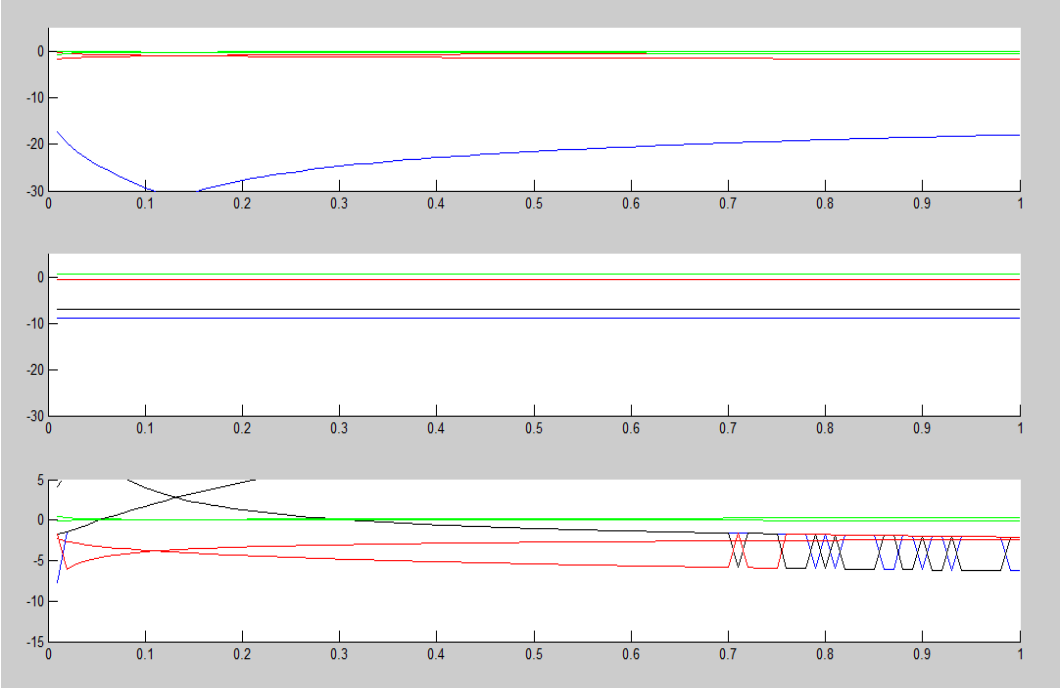
That is the number in line 3 of table 3.4. In the same way, all numbers were shown in Table 3.4, Table 3.5, and Table 3.6.

Appendix II Detailed Data of single parameter traversal analysis

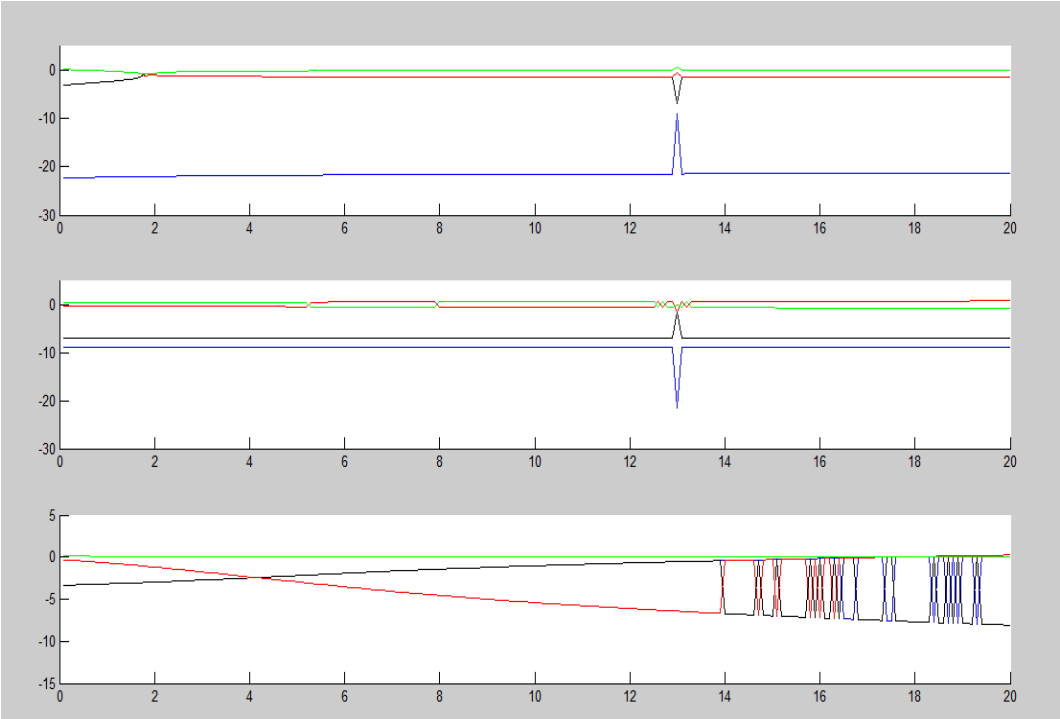
k_1



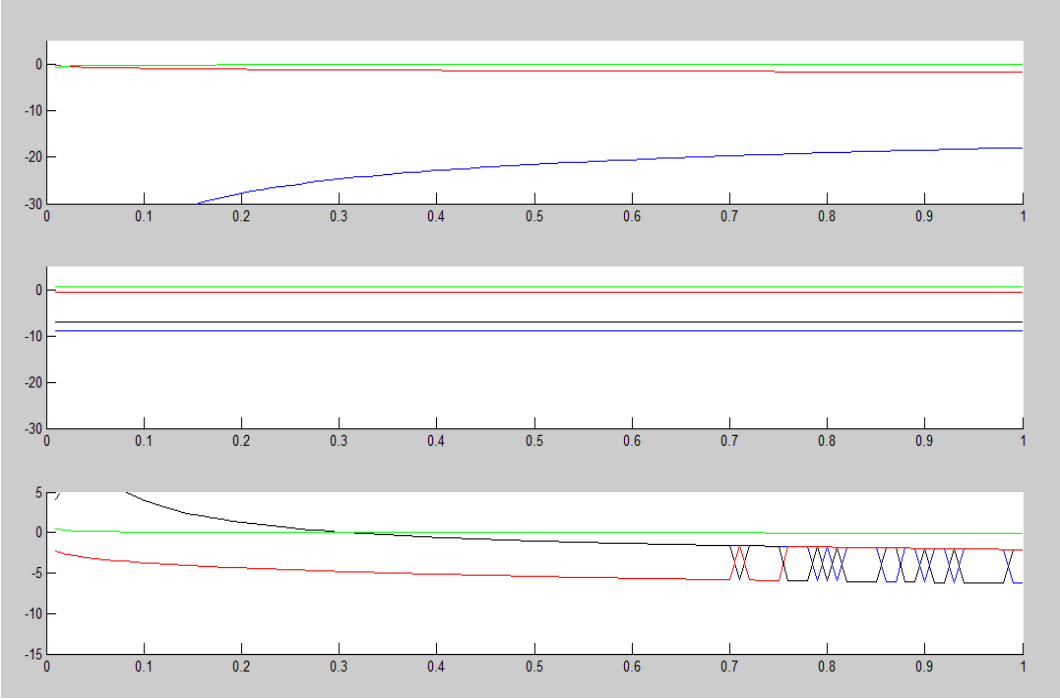
k_1'



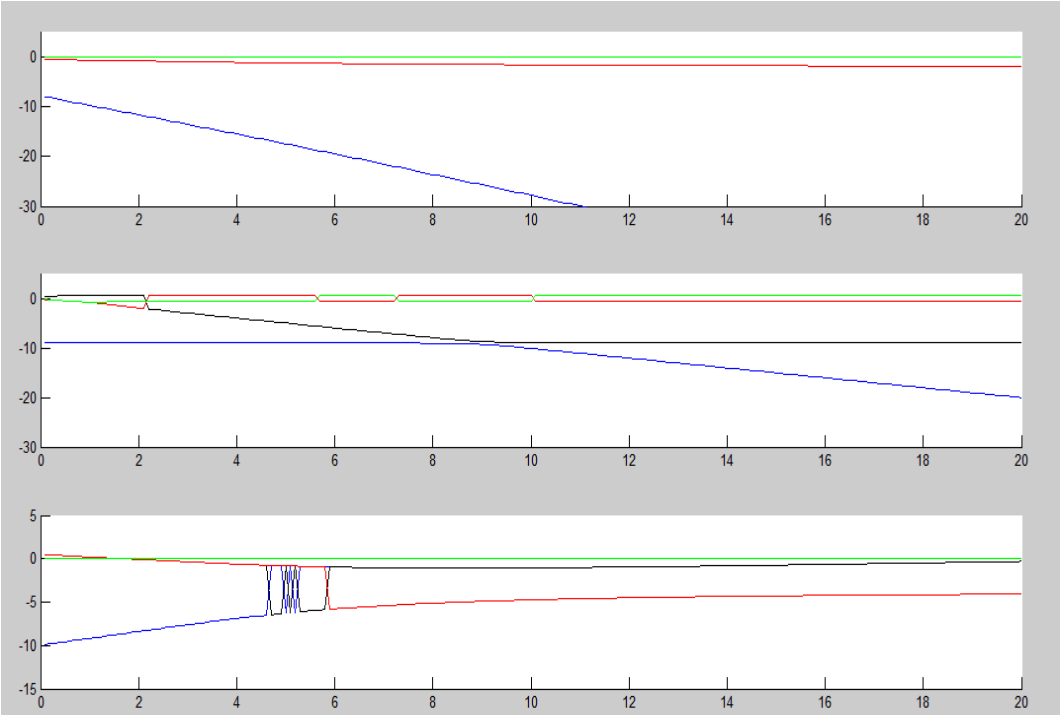
k_2


0
-10
-20
-30
0 2 4 6 8 10 12 14 16 18 20
0
-10
-20
-30
0 2 4 6 8 10 12 14 16 18 20
5
0
-5
-10
-15
0 2 4 6 8 10 12 14 16 18 20

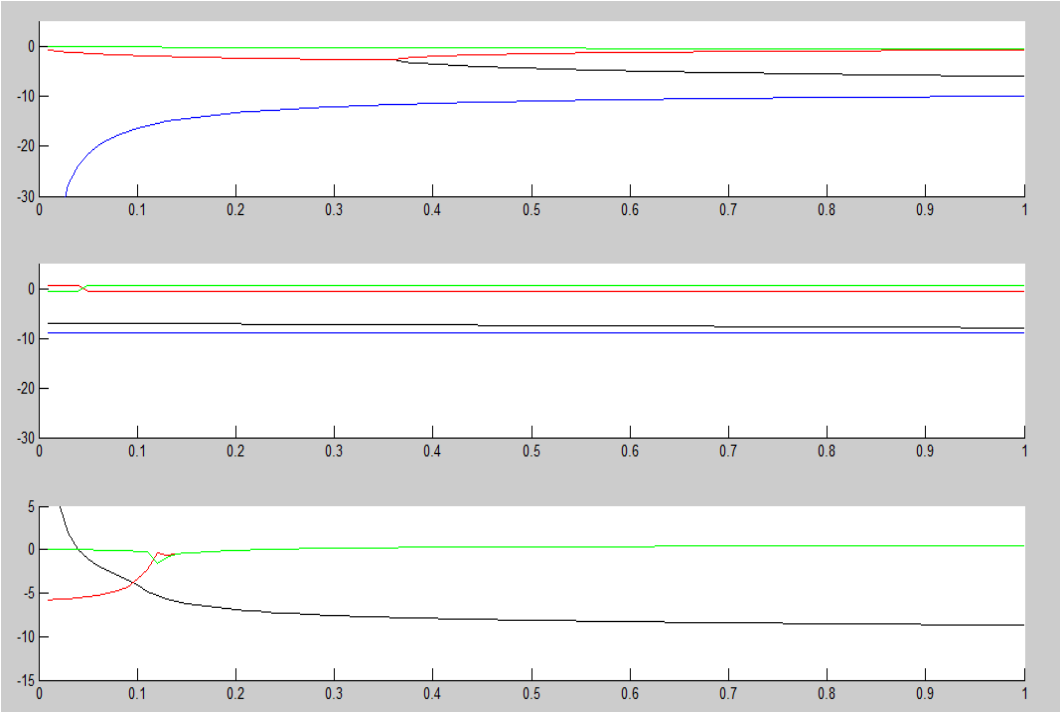
k_2’


0
-10
-20
-30
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0
-10
-20
-30
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
5
0
-5
-10
-15
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

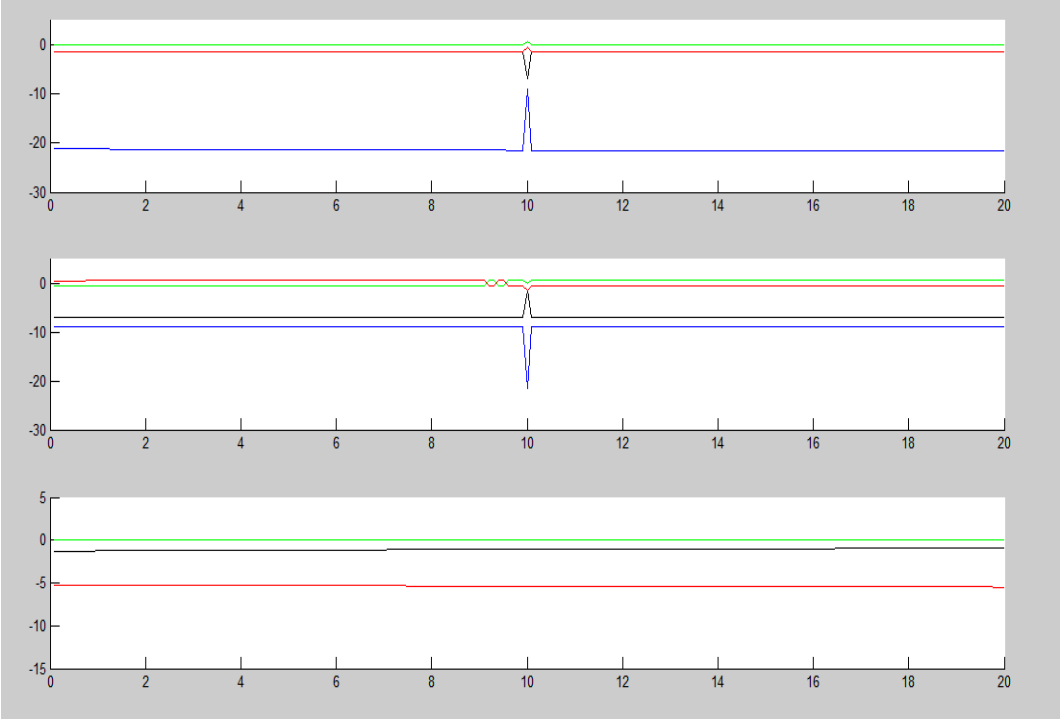
k_3


0
-10
-20
-30
0 2 4 6 8 10 12 14 16 18 20
0
-10
-20
-30
0 2 4 6 8 10 12 14 16 18 20
5
0
-5
-10
-15
0 2 4 6 8 10 12 14 16 18 20

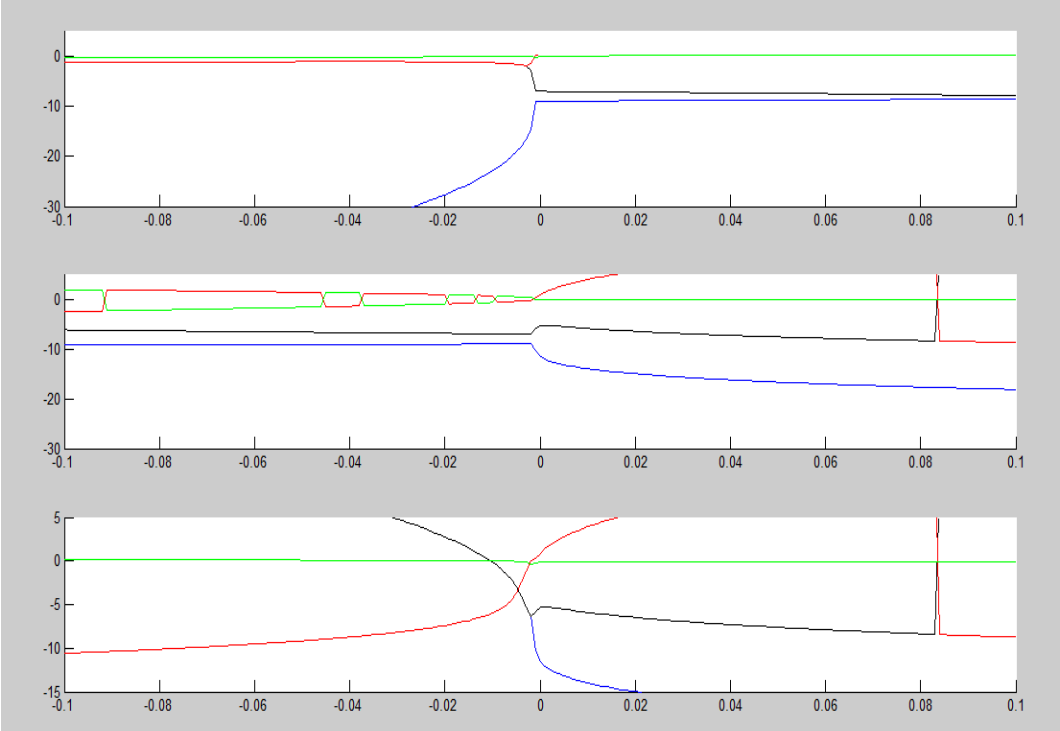
k_3’


0
-10
-20
-30
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0
-10
-20
-30
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
5
0
-5
-10
-15
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

k_4

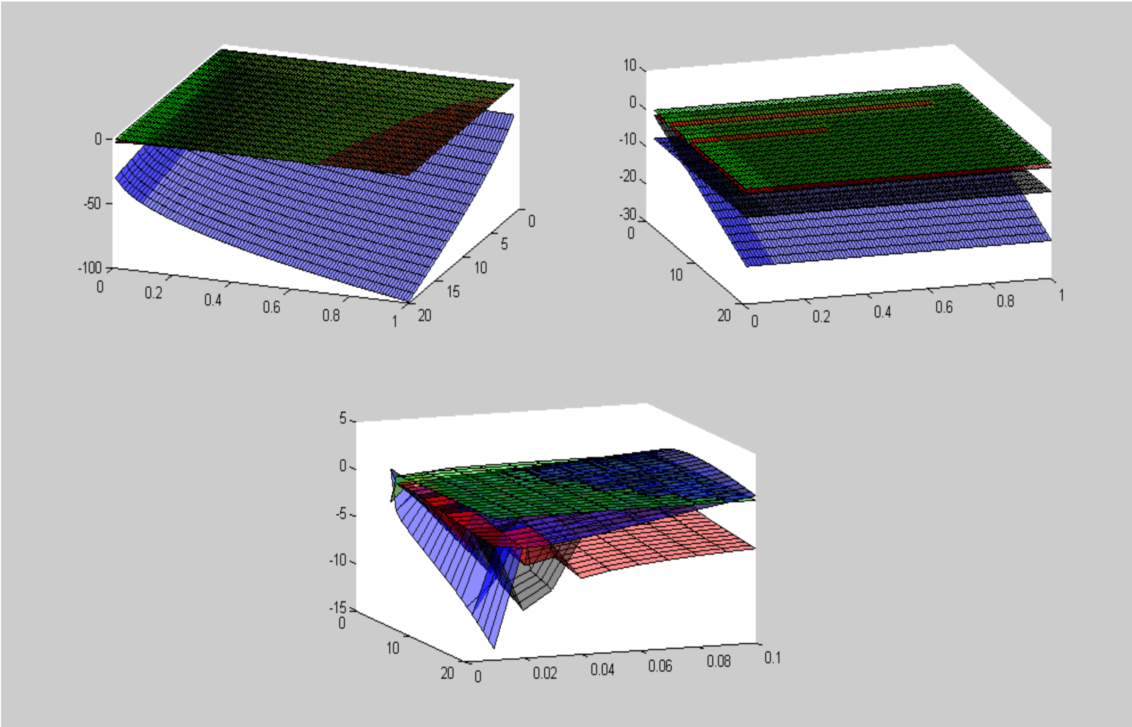

0
-10
-20
-30
0 2 4 6 8 10 12 14 16 18 20
0
-10
-20
-30
0 2 4 6 8 10 12 14 16 18 20
5
0
-5
-10
-15
0 2 4 6 8 10 12 14 16 18 20

k_4’


0
-10
-20
-30
-0.1 -0.08 -0.06 -0.04 -0.02 0 0.02 0.04 0.06 0.08 0.1
0
-10
-20
-30
-0.1 -0.08 -0.06 -0.04 -0.02 0 0.02 0.04 0.06 0.08 0.1
5
0
-5
-10
-15
-0.1 -0.08 -0.06 -0.04 -0.02 0 0.02 0.04 0.06 0.08 0.1

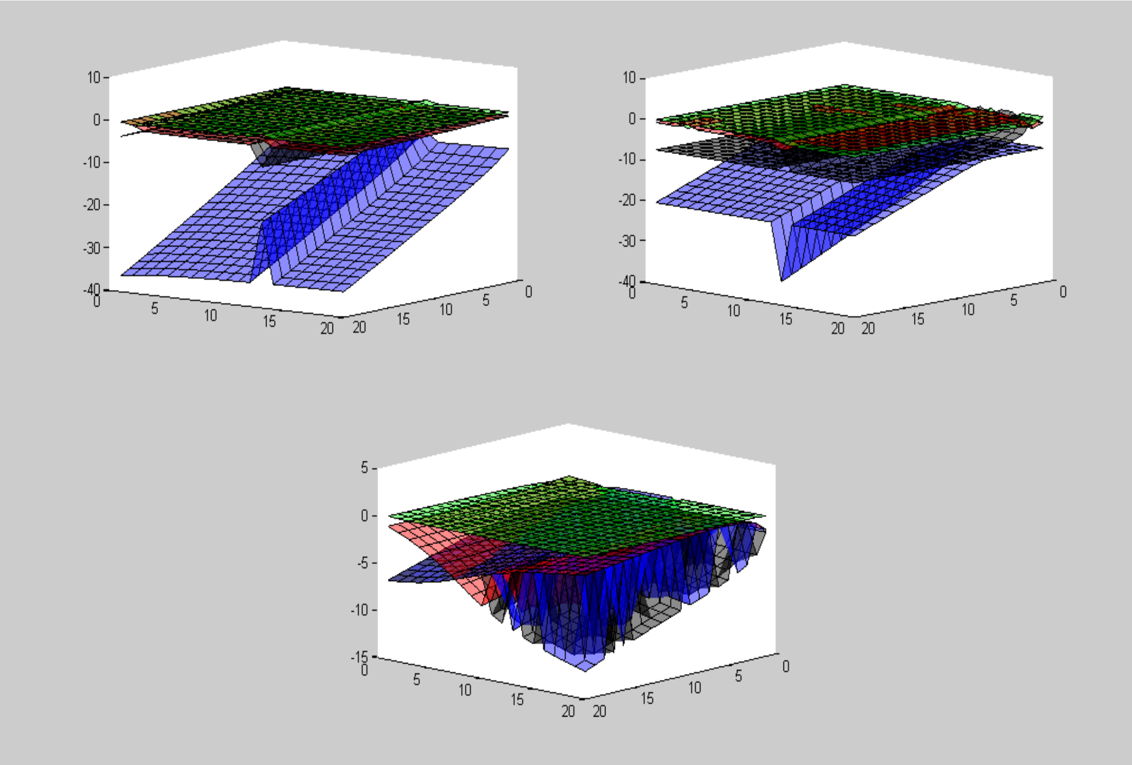
Appendix III Detailed Data of two parameters traversal analysis

k_1&k_1’



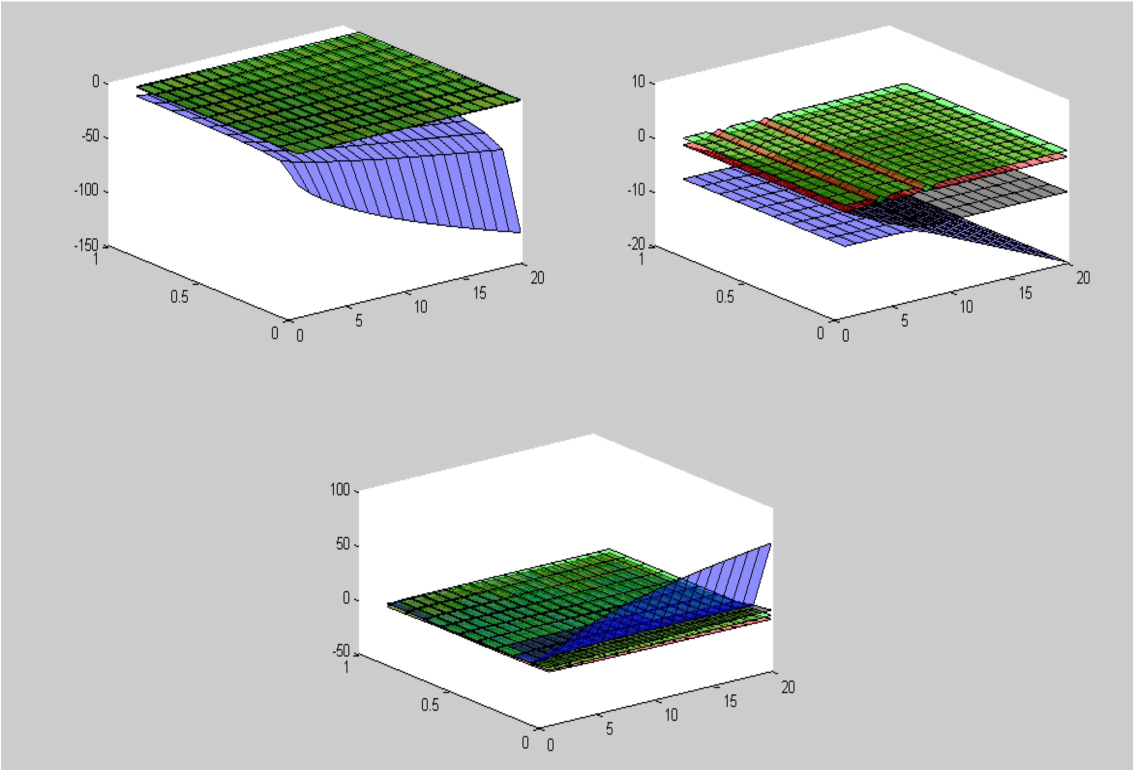
k1’ / k1										
	1	2	2	2	2	2	2	2	2	2
	2	2	2	2	2	2	2	1	1	1
	2	2	2	2	2	1	1	1	1	1
	2	2	2	2	1	1	1	1	1	1
	2	2	2	2	1	1	1	1	1	1
	2	2	2	2	1	1	1	1	1	1
	2	2	2	2	1	1	1	1	1	1
	2	2	2	2	2	1	1	1	1	1
	2	2	2	2	1	1	1	1	1	1
	2	2	2	2	1	1	1	1	1	1
	2	2	2	1	1	1	1	1	1	1
	2	2	2	1	1	1	1	1	1	1
	2	2	2	1	1	1	1	1	1	1
	2	2	1	1	1	1	1	1	1	1
	2	2	1	1	1	1	1	1	1	1
	2	2	1	1	1	1	1	1	1	1
	2	2	1	1	1	1	1	1	1	1
	2	2	1	1	1	1	1	1	1	1
	2	2	1	1	1	1	1	1	1	1
	2	2	1	1	1	1	1	1	1	1

k_1&k_2



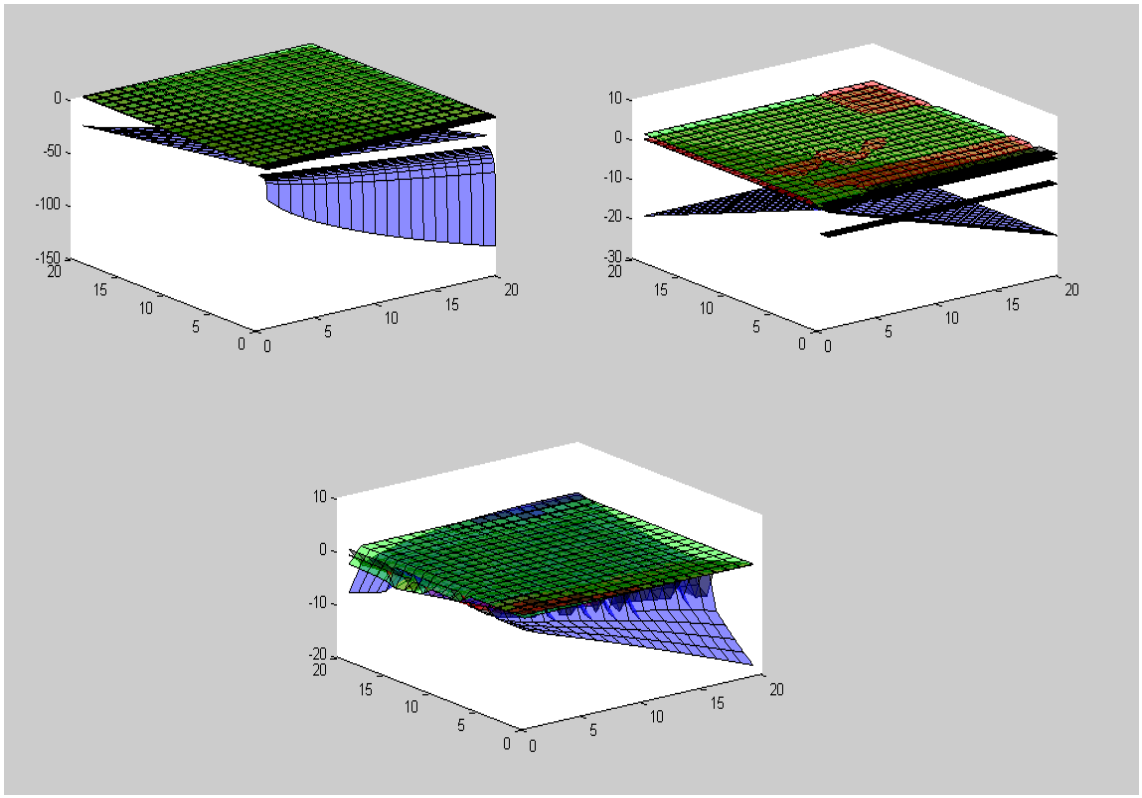
k2 / k1																				
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1
	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1
	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1
	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1
	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1
	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	1
	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2
	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2
	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

k$_1$&k$_2$'



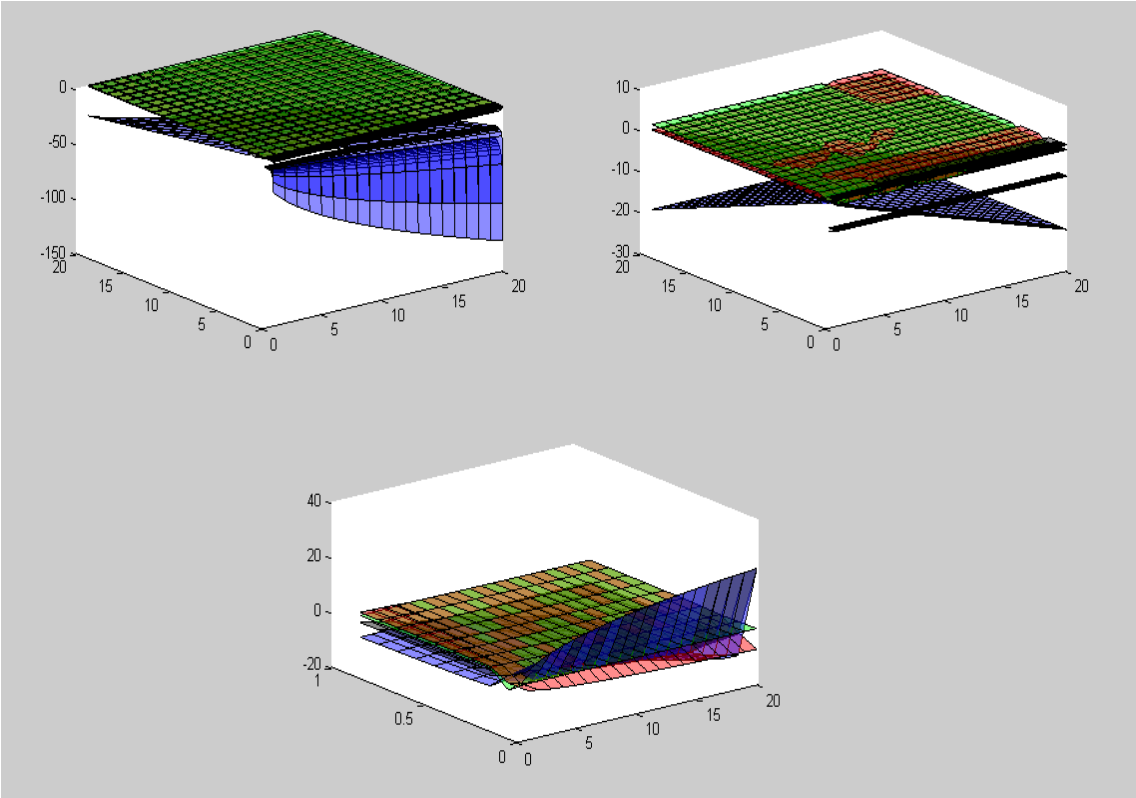
k2' / k1										
	1	1	2	2	2	2	2	2	2	2
	1	1	1	2	2	2	2	2	2	2
	1	1	1	2	2	2	2	2	2	2
	1	1	1	1	2	2	2	2	2	2
	1	1	1	1	2	2	2	2	2	2
	1	1	1	1	2	2	2	2	2	2
	1	1	1	1	2	2	2	2	2	2
	1	1	1	1	2	2	2	2	2	2
	1	1	1	1	2	2	2	2	2	2
	1	1	1	1	2	2	2	2	2	2
	1	1	1	1	2	2	2	2	2	2
	1	1	1	1	1	2	2	2	2	2
	1	1	1	1	1	2	2	2	2	2
	1	1	1	1	1	1	2	2	2	2
	1	1	1	1	1	1	2	2	2	2
	1	1	1	1	1	1	2	2	2	2
	1	1	1	1	1	1	1	2	2	2
	1	1	1	1	1	1	1	2	2	2
	1	1	1	1	1	1	1	1	2	2
	1	1	1	1	1	1	1	1	2	2

k_1&k_3



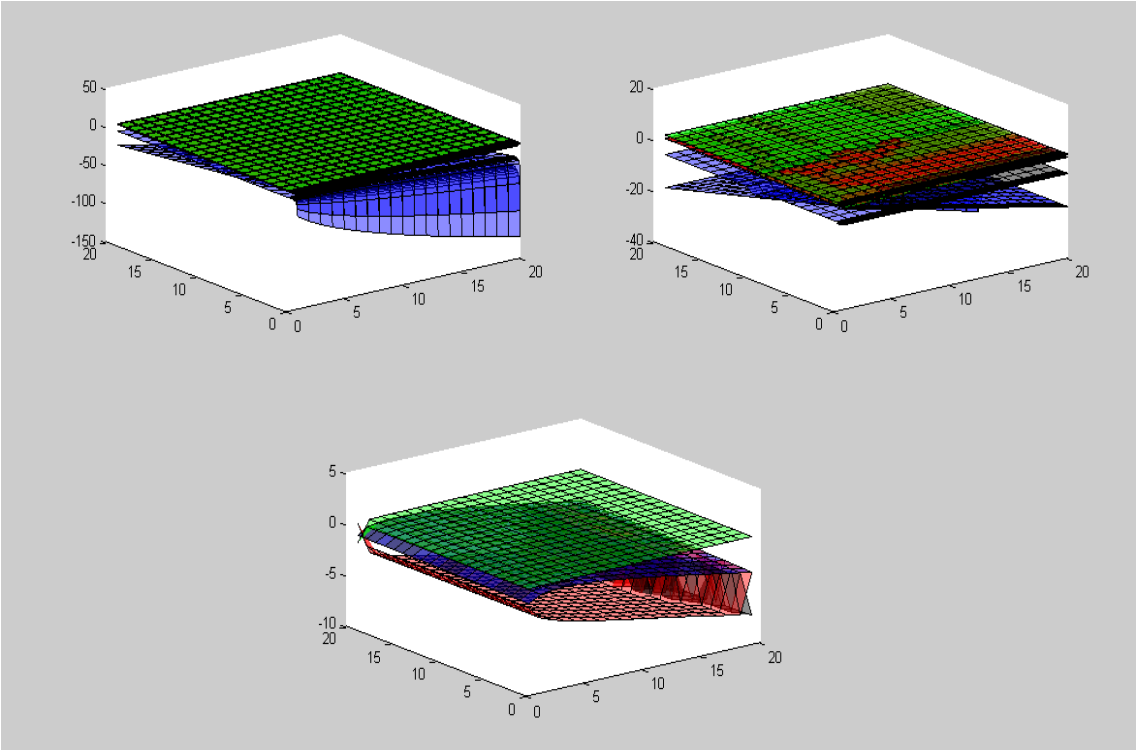
k1 \ k3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
4	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
5	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
8	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
9	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
10	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
11	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1
12	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1
13	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

k_1&k_3’



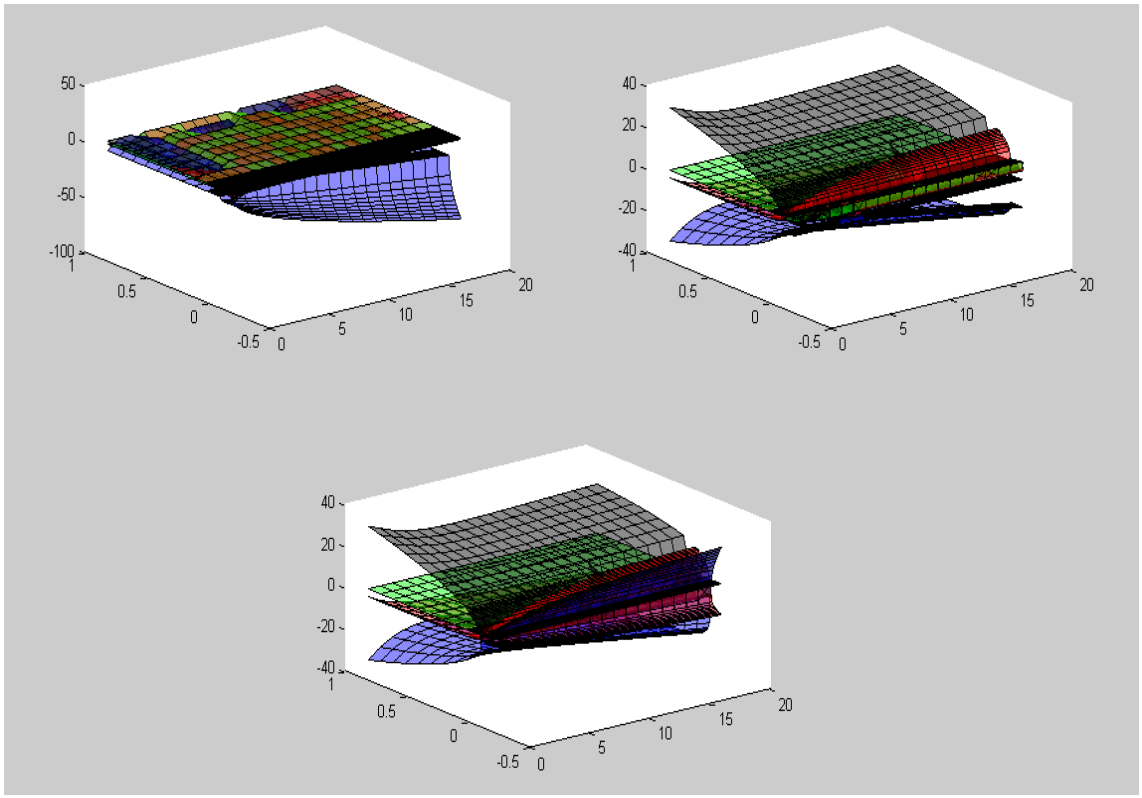
k1 \ k2’	0.01	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1
3	1	2	1	1	1	1	1	1	1	1	1
4	1	2	1	1	1	1	1	1	1	1	1
5	1	2	1	1	1	1	1	1	1	1	1
6	1	2	1	1	1	1	1	1	1	1	1
7	1	2	1	1	1	1	1	1	1	1	1
8	1	2	2	1	1	1	1	1	1	1	1
9	1	2	2	1	1	1	1	1	1	1	1
10	1	2	2	1	1	1	1	1	1	1	1
11	1	2	2	1	1	1	1	1	1	1	1
12	1	2	2	1	1	1	1	1	1	1	1
13	1	2	2	1	1	1	1	1	1	1	1
14	1	2	2	1	1	1	1	1	1	1	1
15	1	2	2	1	1	1	1	1	1	1	1
16	1	2	2	1	1	1	1	1	1	1	1
17	1	2	2	2	1	1	1	1	1	1	1
18	1	2	2	2	1	1	1	1	1	1	1
19	1	2	2	2	1	1	1	1	1	1	1
20	1	2	2	2	1	1	1	1	1	1	1

k_1&k_4



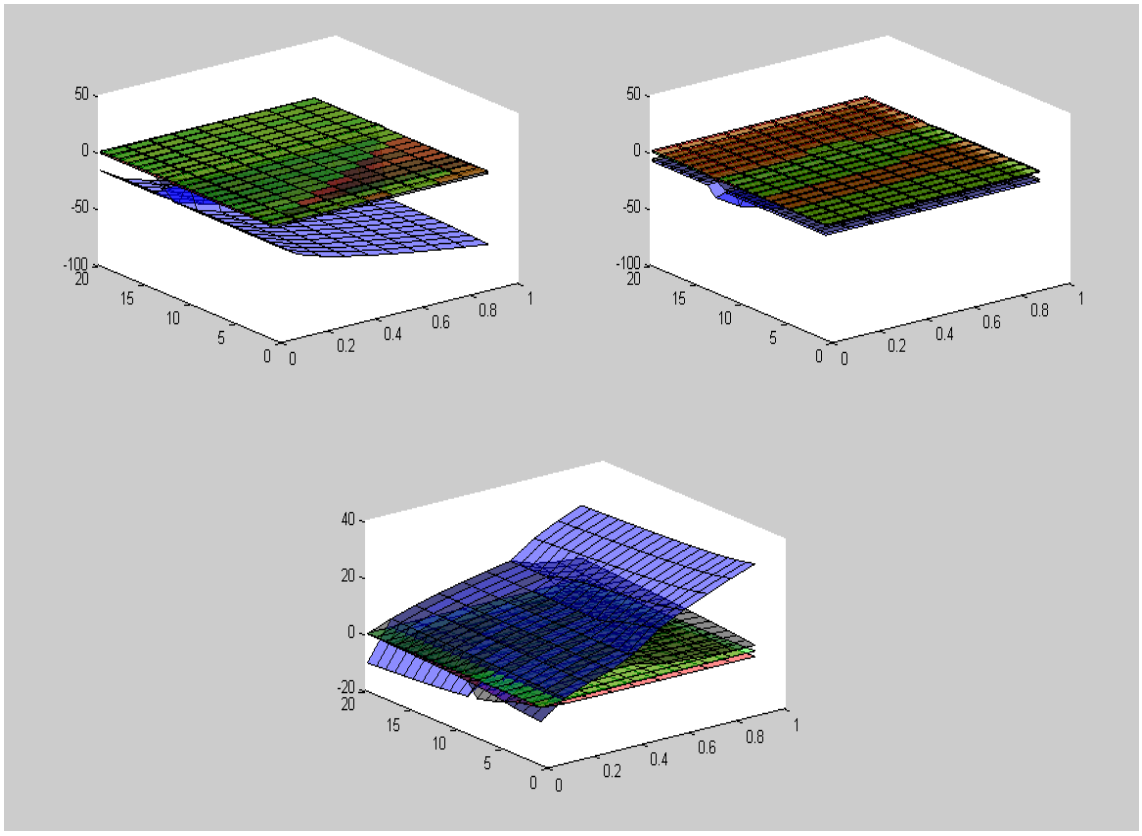
k1 \ k3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
8	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
9	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
10	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
11	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
12	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1
13	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1
14	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1
15	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1
16	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1
17	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1
18	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

k_1&k_4’



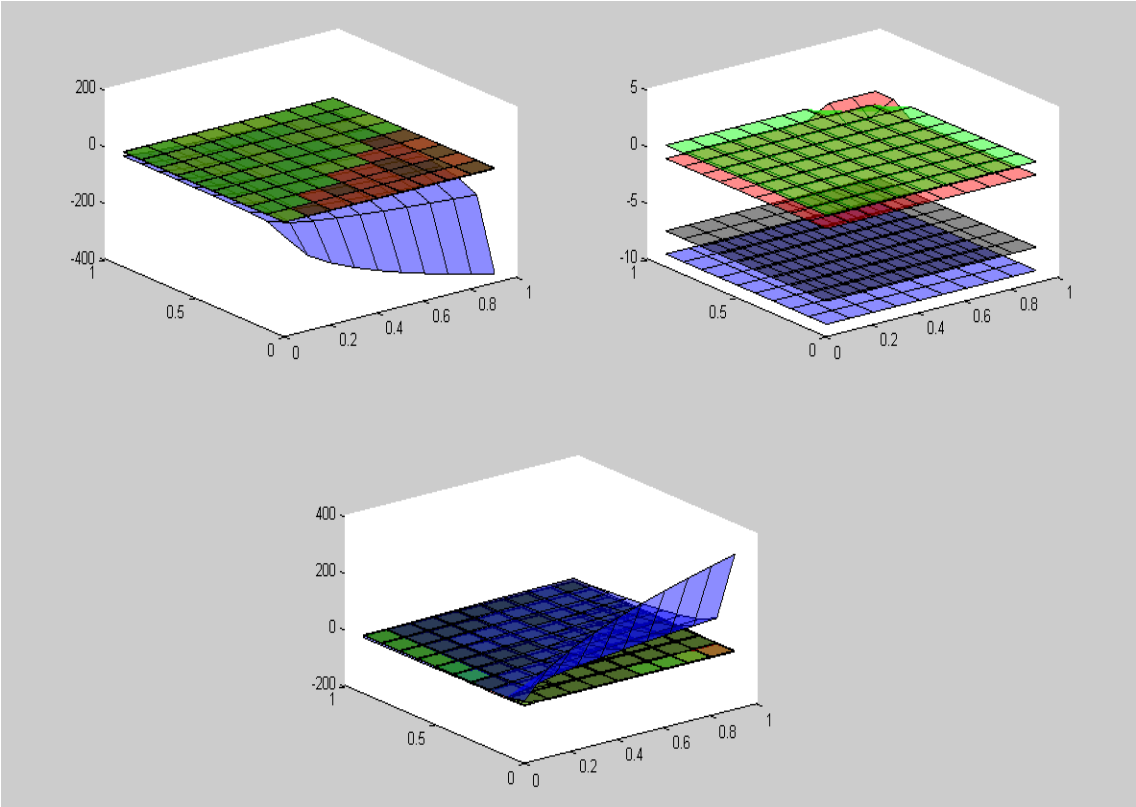
k1 \ k2’	-0.1	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
1	1	1	1	1	1	1	1	1	2	2	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
13	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
14	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
15	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
16	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
17	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
18	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
19	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
20	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0

k_1'&k_2



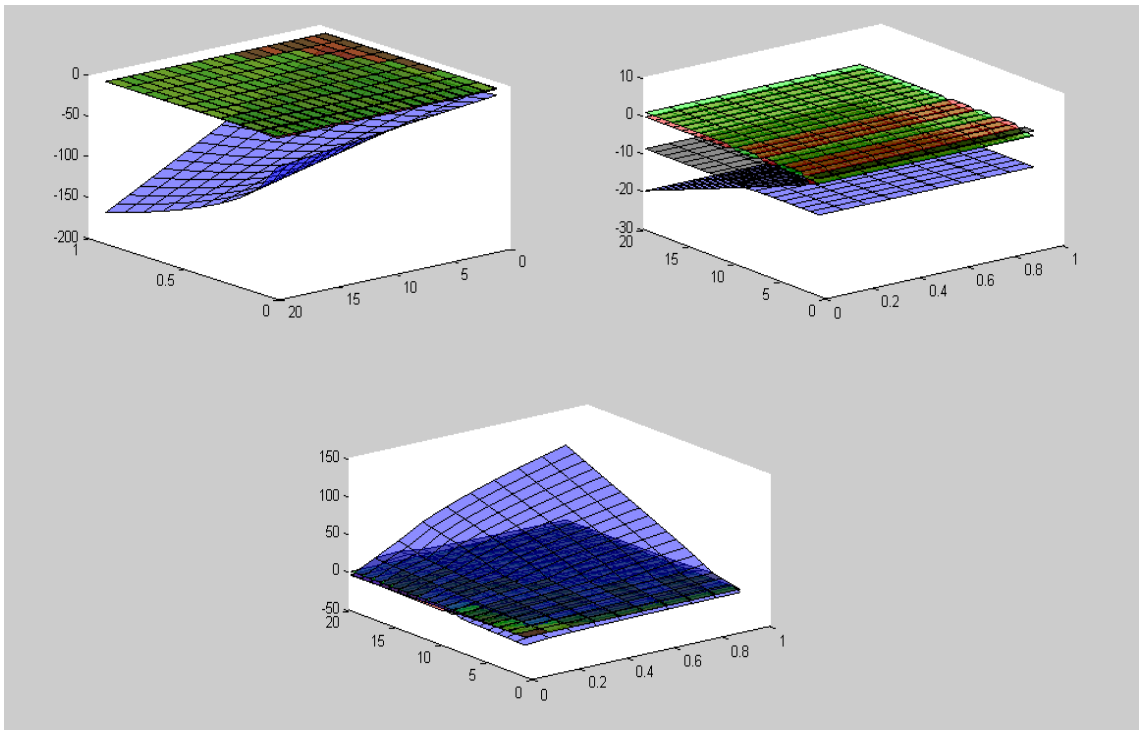
k1' \ k2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.01	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.3	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.4	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.5	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.6	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.7	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.8	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.9	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

k_1'&k_2'



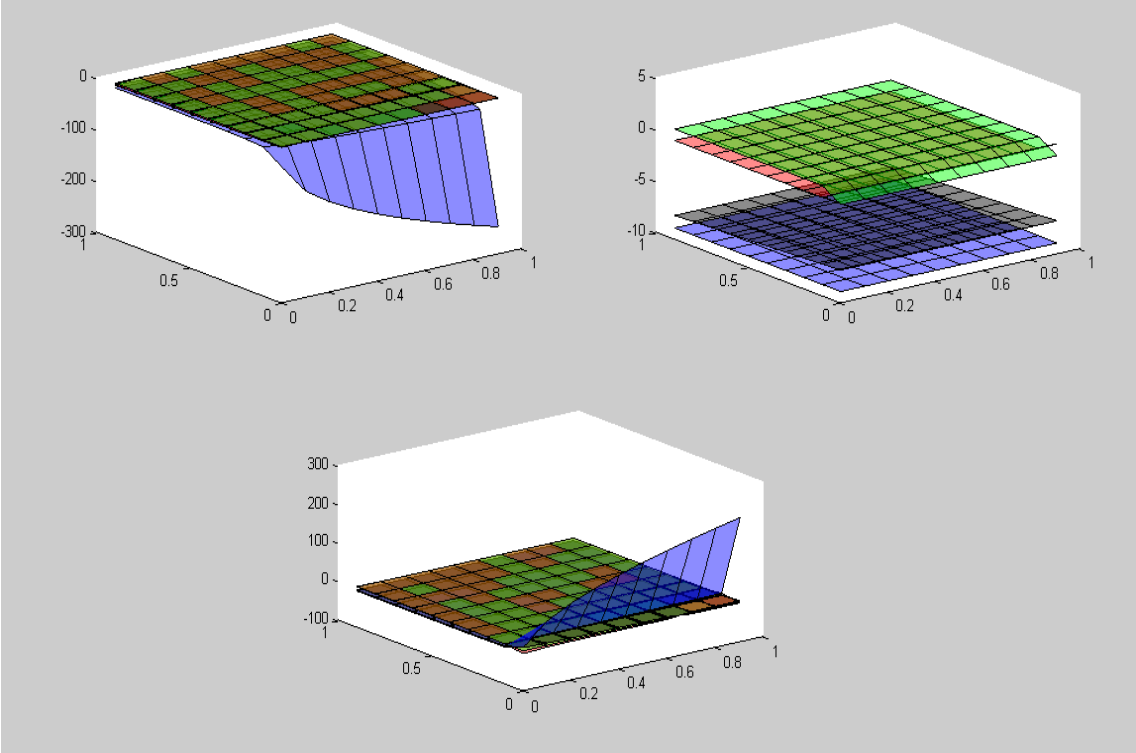
k1' \ k2'	0.01	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.01	1	1	2	2	2	2	2	2	2	2	2
0.1	1	1	1	1	1	1	1	1	1	1	1
0.2	0	1	1	1	1	1	1	1	1	1	1
0.3	0	1	1	1	1	1	1	1	1	1	1
0.4	0	1	1	1	1	1	1	1	1	1	1
0.5	0	1	1	1	1	1	1	1	1	1	1
0.6	0	1	1	1	1	1	1	1	1	1	1
0.7	0	1	1	1	1	1	1	1	1	1	1
0.8	0	1	1	1	1	1	1	1	1	1	1
0.9	0	1	1	1	1	1	1	1	1	1	1
1	0	1	1	1	1	1	1	1	1	1	1

k_1'&k_3



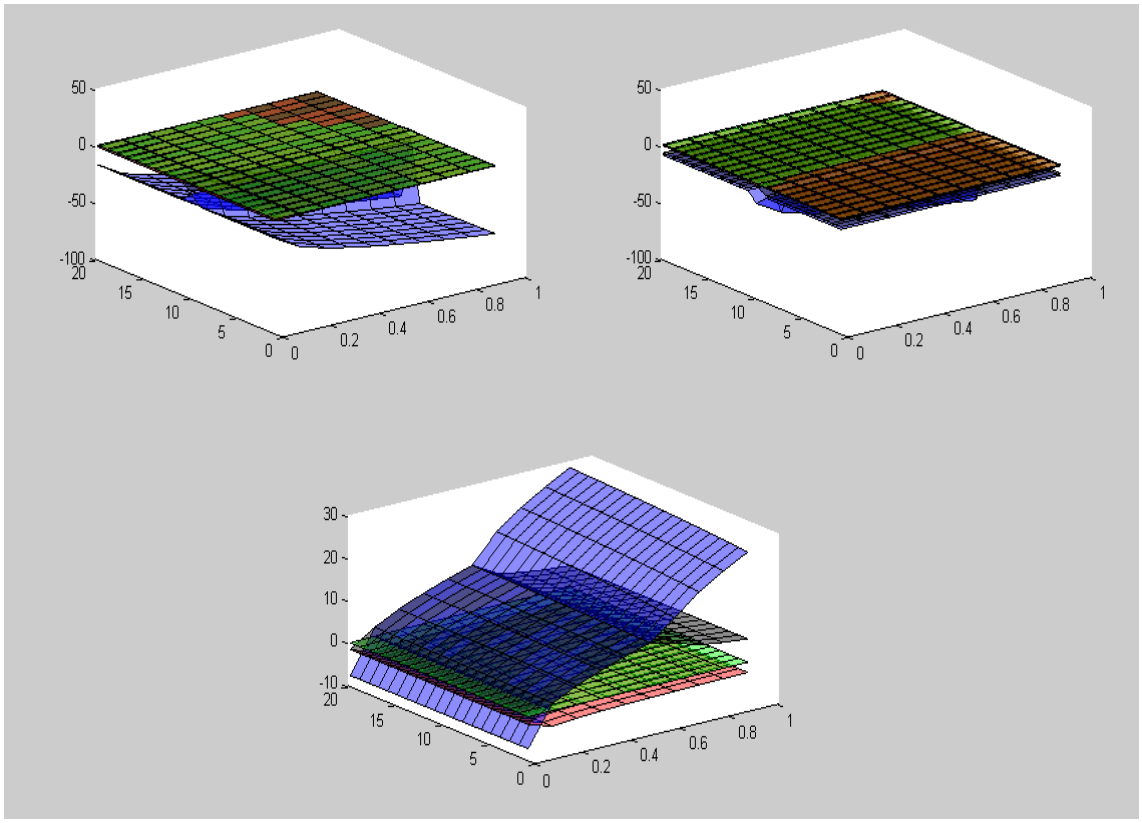
k1' \ k3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.01	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

k_1'&k_3'



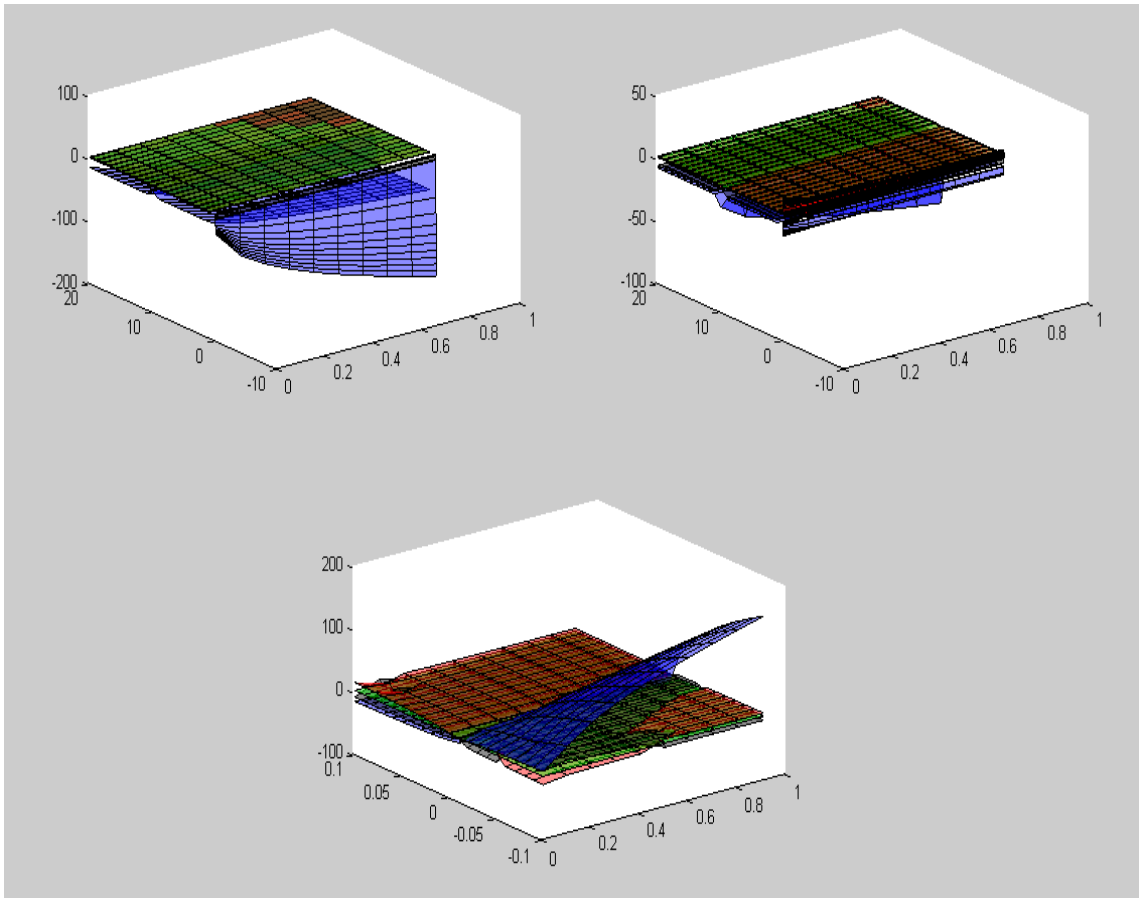
k1' \ k3'	0.01	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.01	1	2	1	1	1	1	1	1	1	1	1
0.1	1	2	2	2	2	1	1	1	1	1	1
0.2	1	2	2	2	2	2	2	1	1	1	1
0.3	1	1	2	2	2	2	2	2	2	1	1
0.4	1	1	2	2	2	2	2	2	2	2	2
0.5	1	1	2	2	2	2	2	2	2	2	2
0.6	1	1	2	2	2	2	2	2	2	2	2
0.7	1	1	2	2	2	2	2	2	2	2	2
0.8	1	1	1	2	2	2	2	2	2	2	2
0.9	1	1	1	2	2	2	2	2	2	2	2
1	1	1	1	2	2	2	2	2	2	2	2

k_1'&k_4



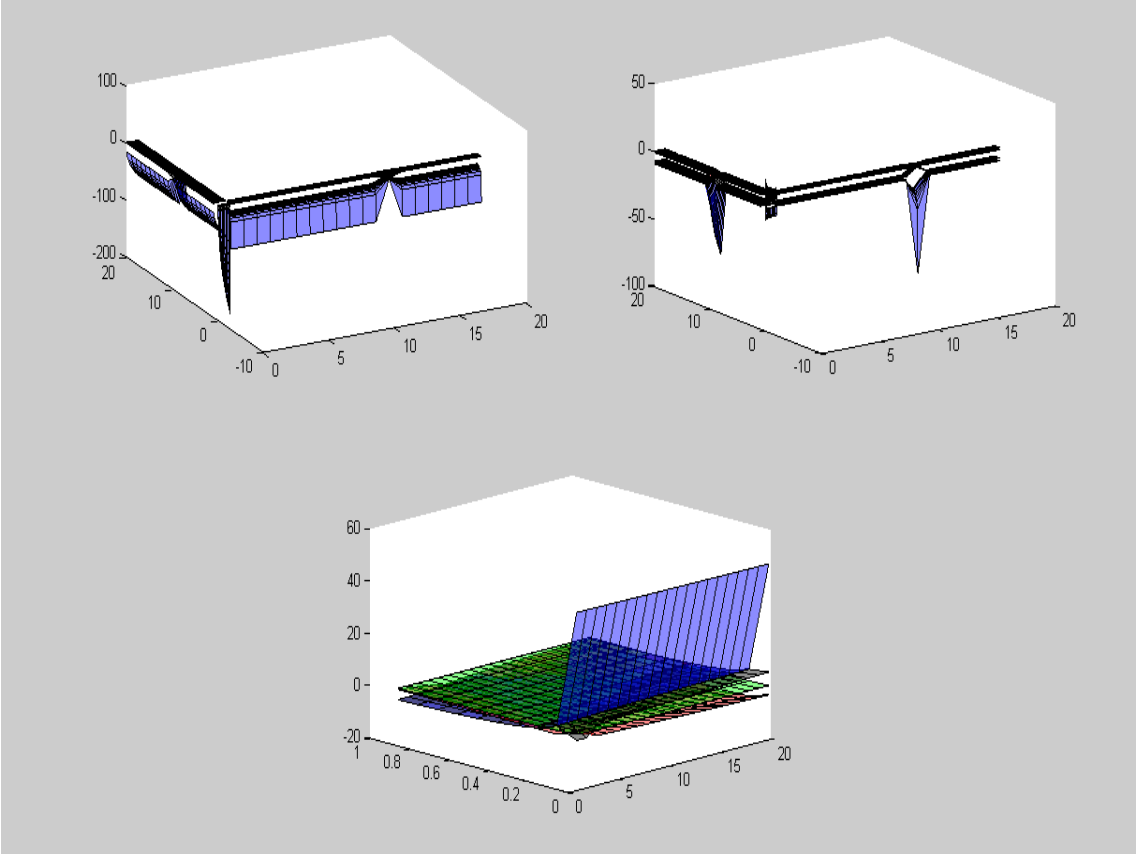
k1' \ k4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.01	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

k_1'&k_4'



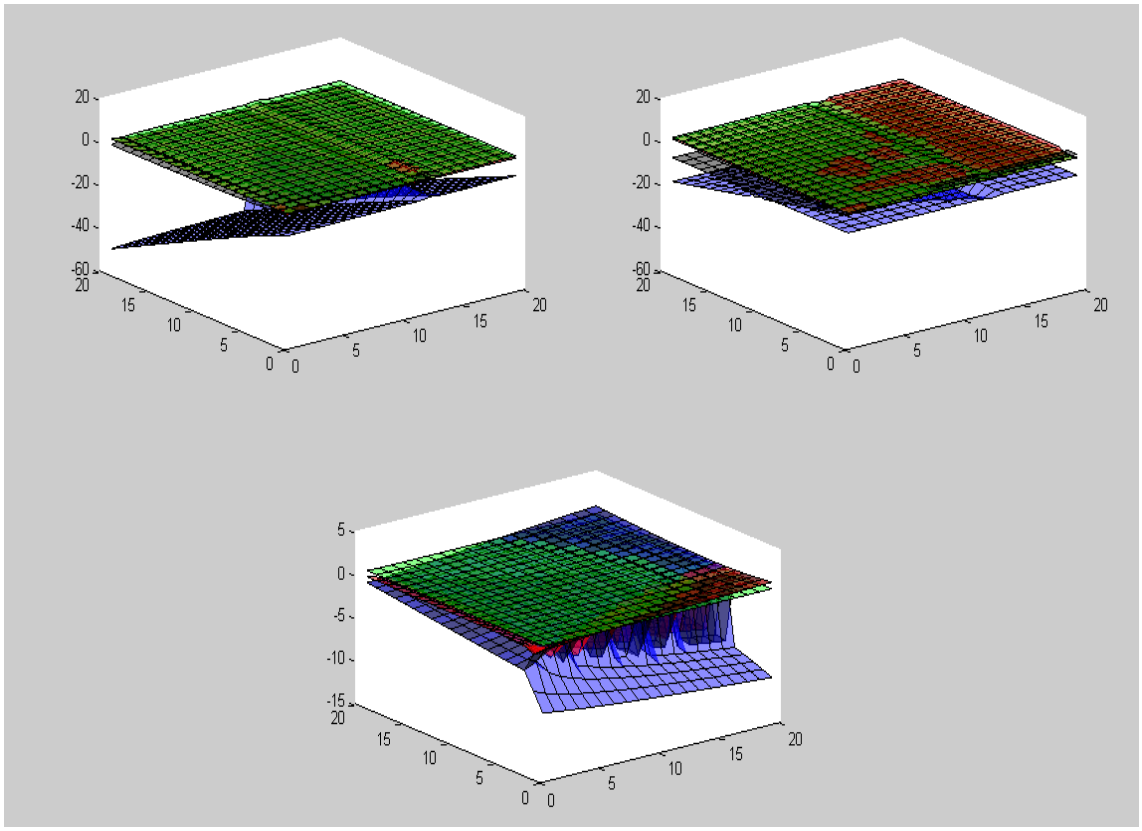
k1' \ k4'	-0.1	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
0.01	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
0.1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.2	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.3	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.4	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.5	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.6	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.7	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.8	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.9	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0

k_2&k_2’



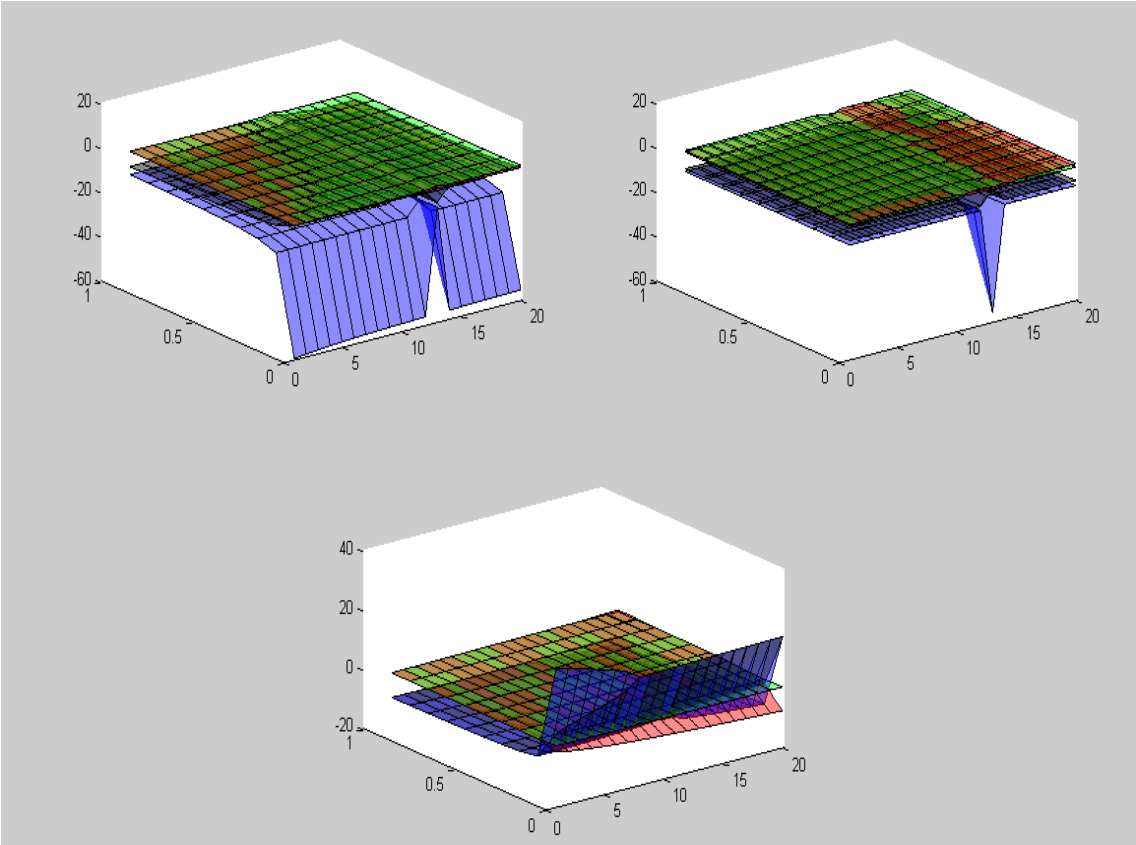
k2 \ k2’	0.01	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	0	1	1	1	1	1	1	2	2	2	2
2	0	1	1	1	1	1	2	2	2	2	2
3	0	1	1	1	1	1	2	2	2	2	2
4	0	1	1	1	1	2	2	2	2	2	2
5	1	1	1	1	1	2	2	2	2	2	2
6	1	1	1	1	2	2	2	2	2	2	2
7	1	1	1	1	2	2	2	2	2	2	2
8	1	1	1	1	2	2	2	2	2	2	2
9	1	1	1	1	2	2	2	2	2	2	2
10	1	1	1	1	2	2	2	2	2	2	2
11	1	1	1	1	2	2	2	2	2	2	2
12	1	1	1	1	2	2	2	2	2	2	2
13	1	1	1	1	2	2	2	2	2	2	2
14	1	1	1	1	2	2	2	2	2	2	2
15	1	1	1	1	1	2	2	2	2	2	2
16	1	1	1	1	1	2	2	2	2	2	2
17	1	1	1	1	1	2	2	2	2	2	2
18	1	1	1	1	1	1	2	2	2	2	2
19	1	1	1	1	1	1	2	2	2	2	2
20	1	1	1	1	1	1	2	2	2	2	2

k_2&k_3



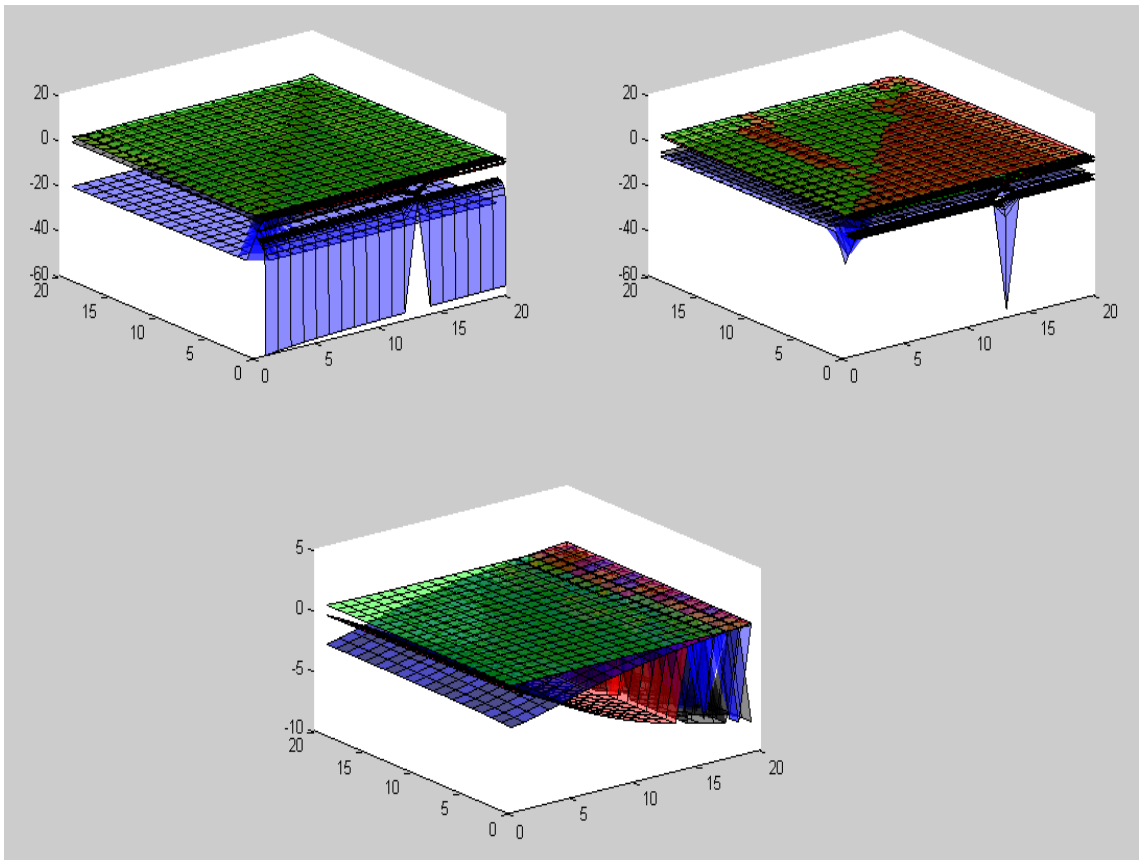
k2 \ k3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
8	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
9	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
10	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
11	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
12	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
13	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1
14	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1
15	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1
16	1	1	1	1	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1
17	1	1	1	1	1	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

k_2&k_3’



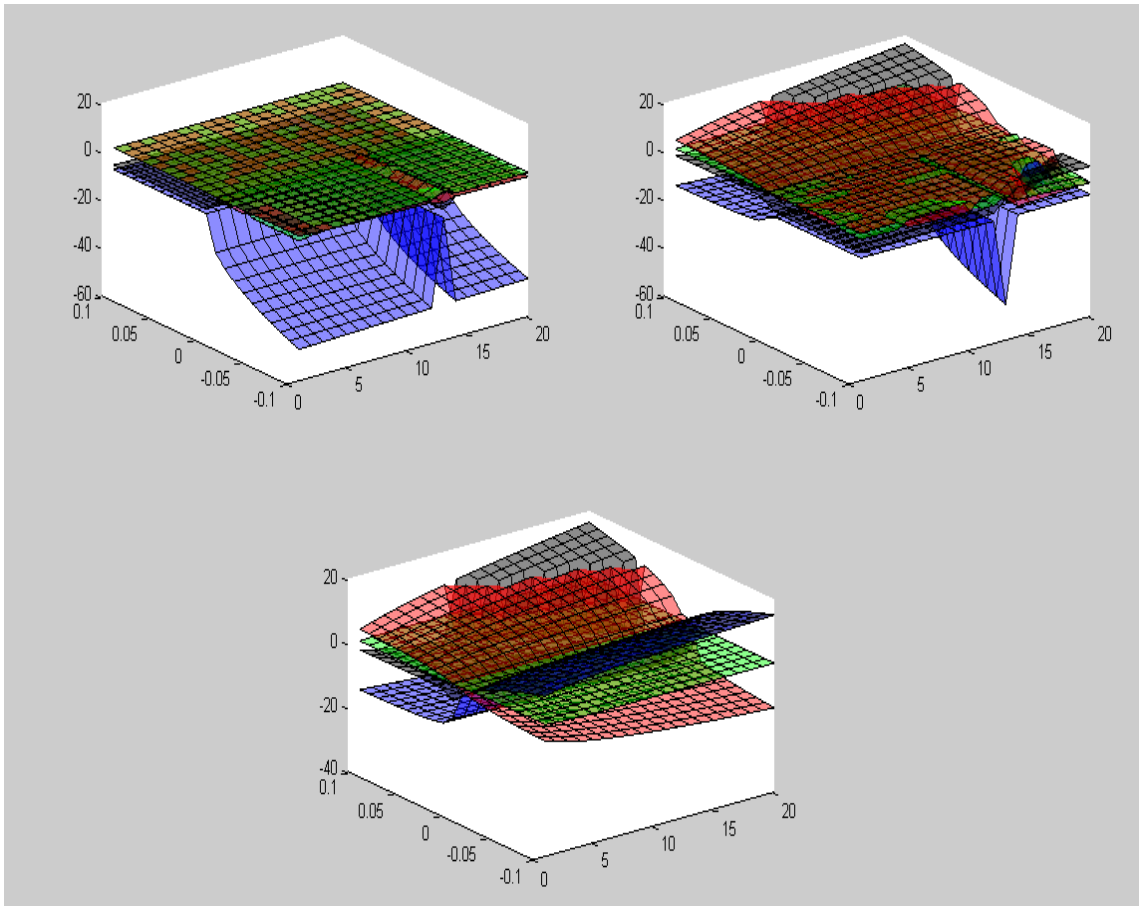
k2 \ k3’	0.01	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	1	2	2	1	1	1	1	1	1	1	1
2	1	2	2	1	1	1	1	1	1	1	1
3	1	2	2	1	1	1	1	1	1	1	1
4	1	2	2	1	1	1	1	1	1	1	1
5	1	2	2	1	1	1	1	1	1	1	1
6	1	2	2	1	1	1	1	1	1	1	1
7	1	2	2	1	1	1	1	1	1	1	1
8	1	2	2	1	1	1	1	1	1	1	1
9	1	2	2	1	1	1	1	1	1	1	1
10	1	2	2	1	1	1	1	1	1	1	1
11	1	2	2	1	1	1	1	1	1	1	1
12	1	2	2	1	1	1	1	1	1	1	1
13	1	2	2	1	1	1	1	1	1	1	1
14	1	2	2	1	1	1	1	1	1	1	1
15	1	2	2	1	1	1	1	1	1	1	1
16	1	2	2	1	1	1	1	1	1	1	1
17	1	2	2	1	1	1	1	1	1	1	1
18	1	2	2	1	1	1	1	1	1	1	1
19	1	2	2	1	1	1	1	1	1	1	1
20	1	2	2	1	1	1	1	1	1	1	1

k_2&k_4



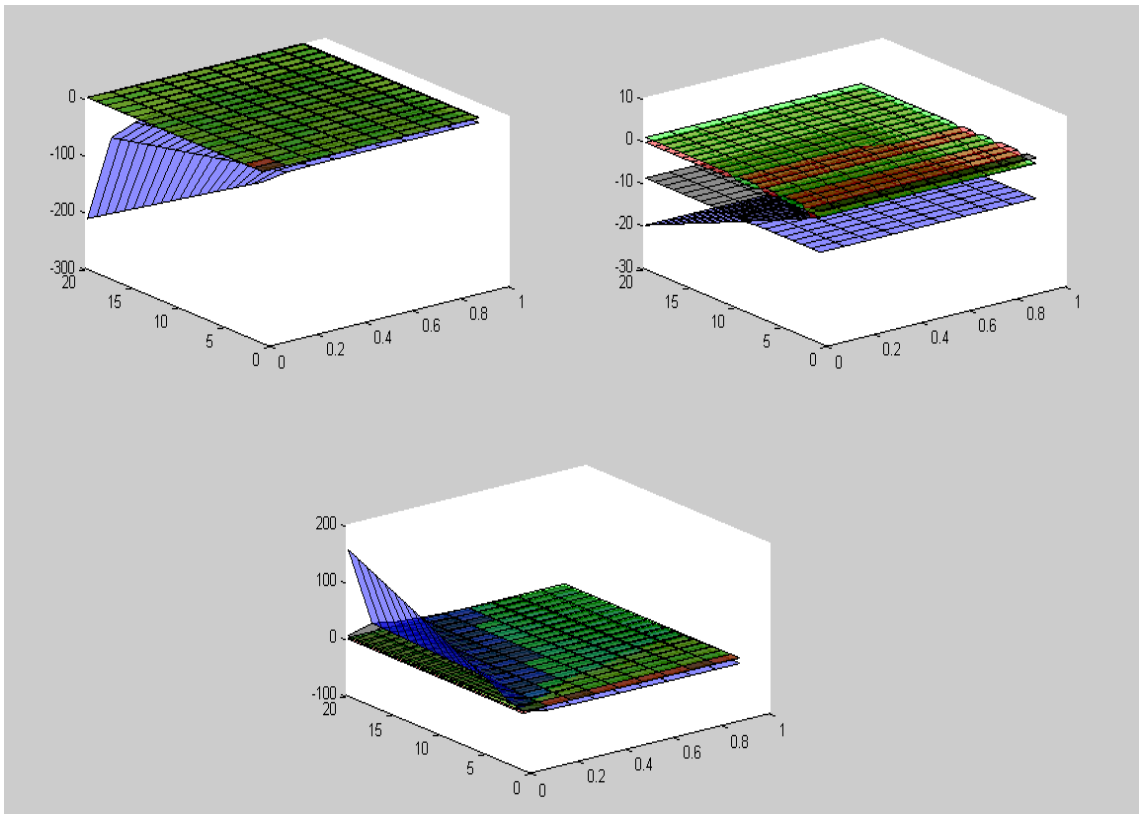
k2 \ k4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1
3	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1
4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1
5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1
6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
8	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
9	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
10	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
11	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
12	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
13	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
14	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
15	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
16	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
17	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1
18	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

k_2&k_4’



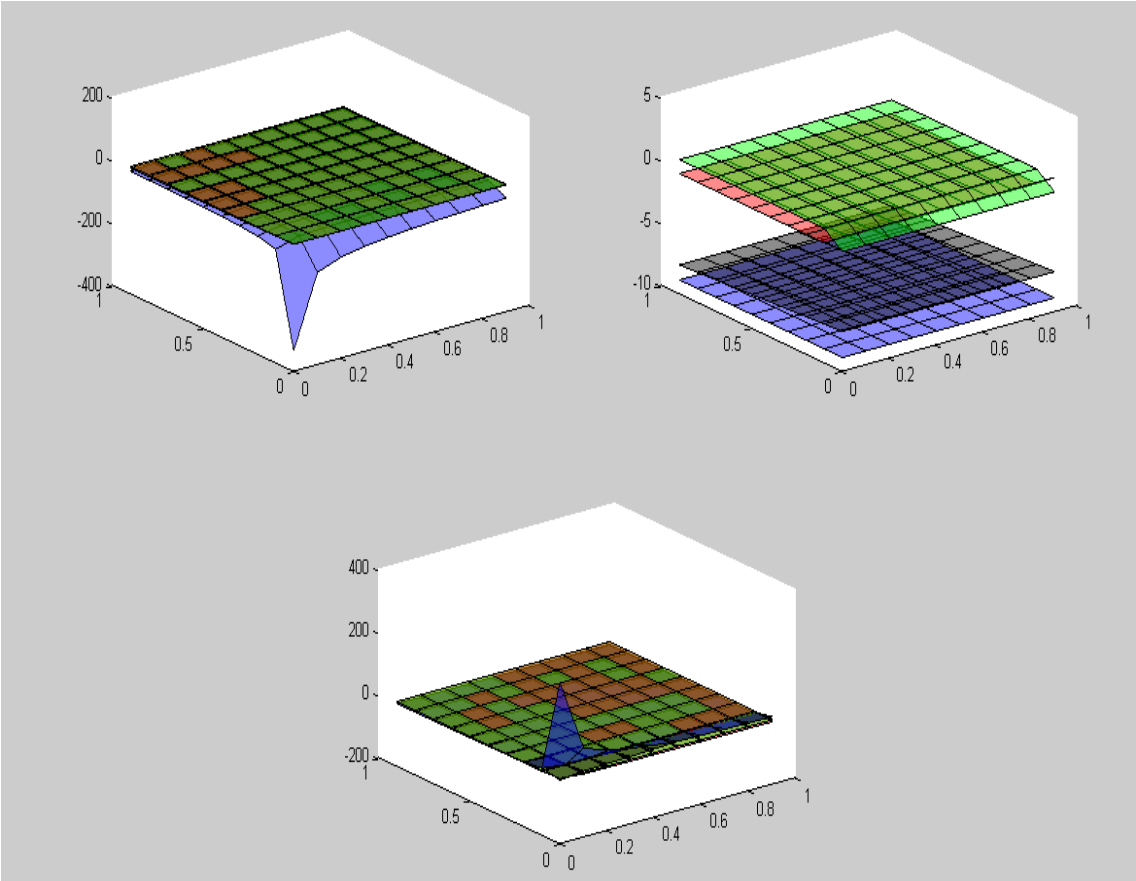
k2 \ k4’	-0.1	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
13	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
14	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
15	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
16	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
17	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
18	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
19	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
20	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0

k_2'&k_3



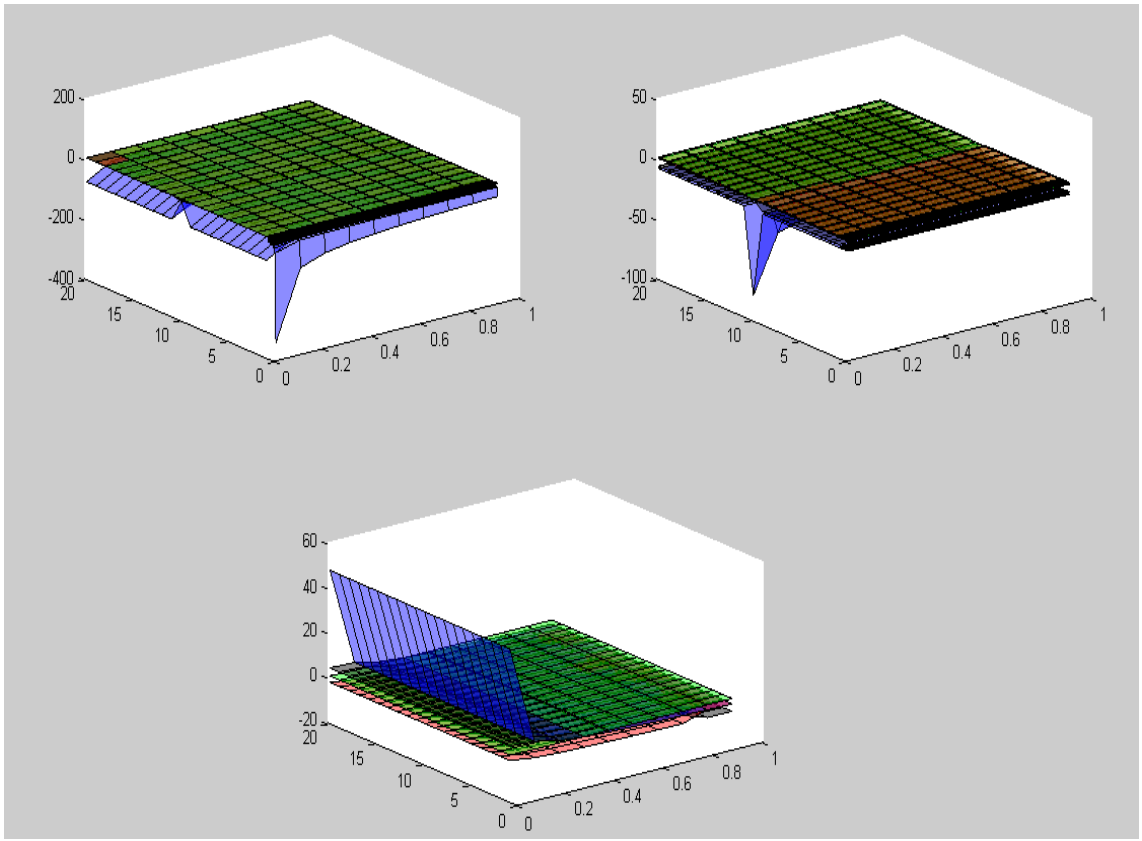
k2' \ k3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.01	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.4	1	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1
0.5	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.6	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.7	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.8	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.9	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

k_2'&k_3'



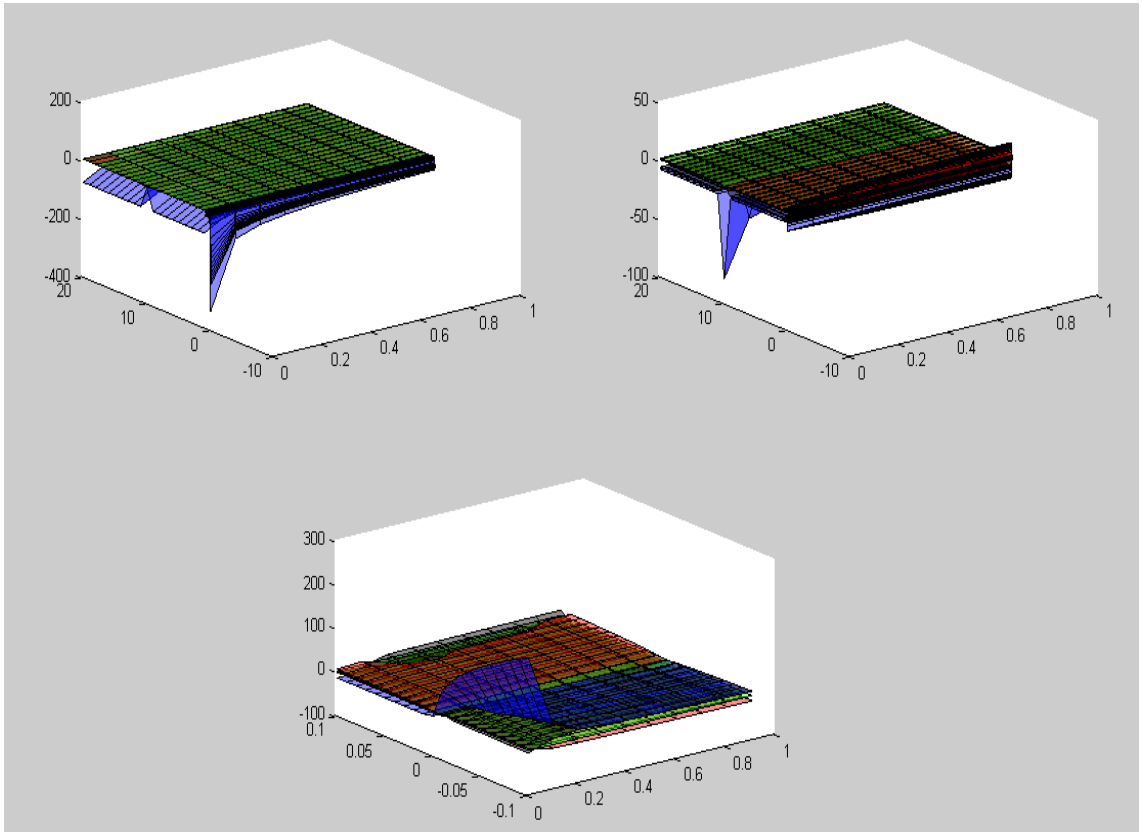
k2' \ k3'	0.01	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.01	1	1	1	2	2	2	2	1	1	1	1
0.1	1	2	2	2	2	2	1	1	1	1	1
0.2	1	2	2	2	1	1	1	1	1	1	1
0.3	1	2	2	1	1	1	1	1	1	1	1
0.4	1	2	2	1	1	1	1	1	1	1	1
0.5	1	2	2	1	1	1	1	1	1	1	1
0.6	1	2	1	1	1	1	1	1	1	1	1
0.7	1	2	1	1	1	1	1	1	1	1	1
0.8	1	2	1	1	1	1	1	1	1	1	1
0.9	1	2	1	1	1	1	1	1	1	1	1
1	1	2	1	1	1	1	1	1	1	1	1

k_2'&k_4



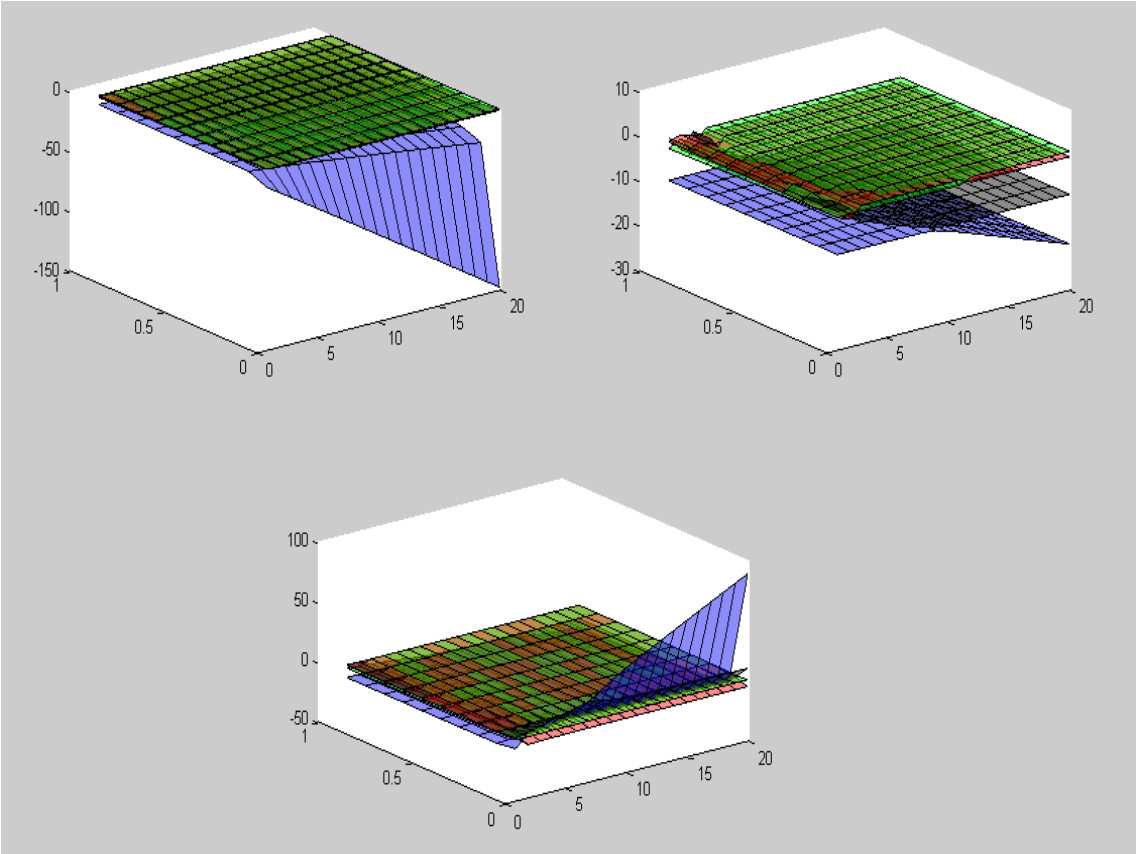
k2' \ k4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.01	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.3	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1
0.4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.7	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.8	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.9	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

k_2'&k_4'



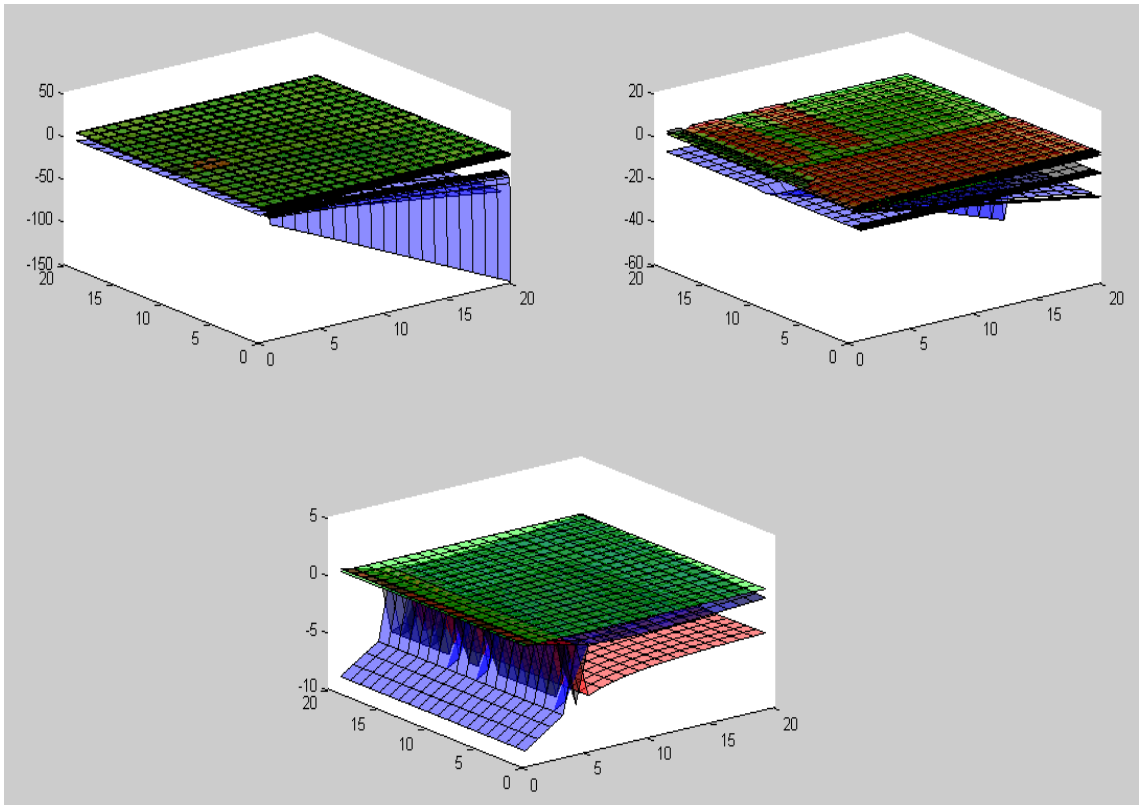
k2' \ k4'	-0.1	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
0.01	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.2	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.3	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.4	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.5	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
0.6	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
0.7	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
0.8	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
0.9	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0

k_3&k_3’



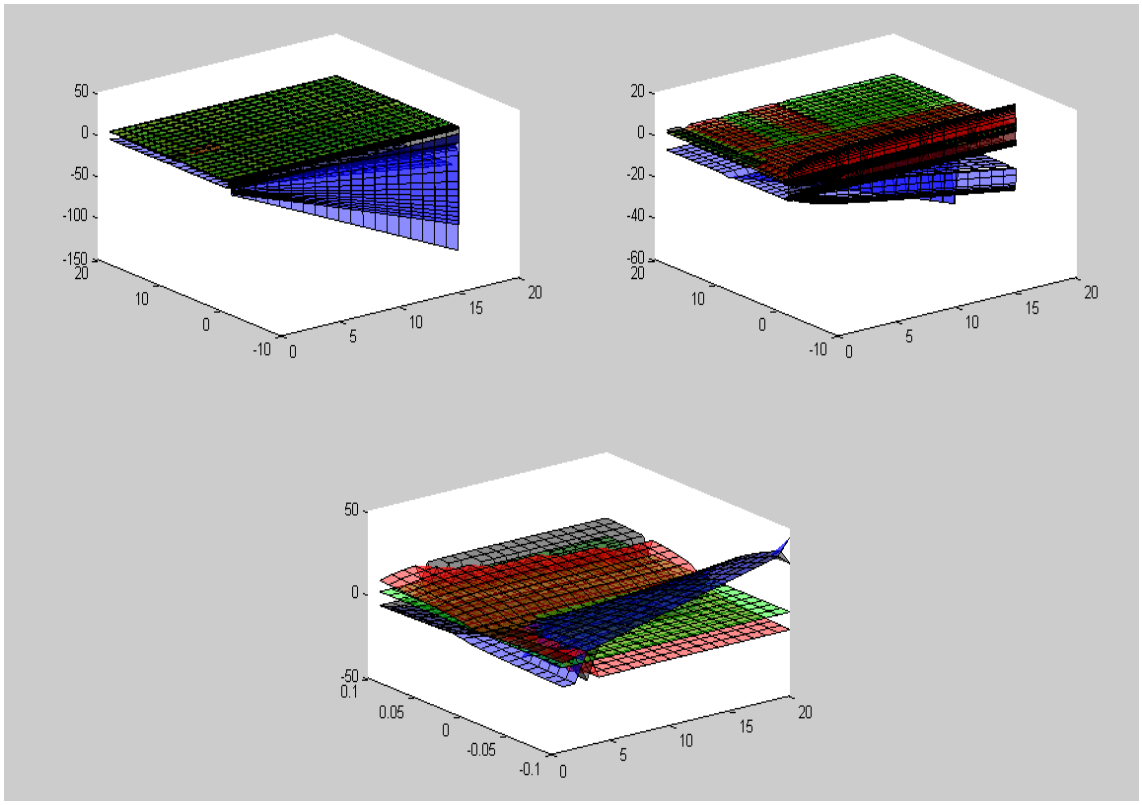
k3 \ k3’	0.01	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	1	1	1	1	1	1	1	1	1	1	1
2	1	2	1	1	1	1	1	1	1	1	1
3	1	2	1	1	1	1	1	1	1	1	1
4	1	2	2	1	1	1	1	1	1	1	1
5	1	2	2	1	1	1	1	1	1	1	1
6	1	2	2	1	1	1	1	1	1	1	1
7	1	2	2	1	1	1	1	1	1	1	1
8	1	2	2	1	1	1	1	1	1	1	1
9	1	2	2	1	1	1	1	1	1	1	1
10	1	2	2	1	1	1	1	1	1	1	1
11	1	2	2	1	1	1	1	1	1	1	1
12	1	2	2	1	1	1	1	1	1	1	1
13	1	2	2	1	1	1	1	1	1	1	1
14	1	2	2	1	1	1	1	1	1	1	1
15	1	2	2	1	1	1	1	1	1	1	1
16	1	2	2	1	1	1	1	1	1	1	1
17	1	2	2	1	1	1	1	1	1	1	1
18	1	2	2	1	1	1	1	1	1	1	1
19	1	2	2	1	1	1	1	1	1	1	1
20	1	2	2	1	1	1	1	1	1	1	1

k_3&k_4



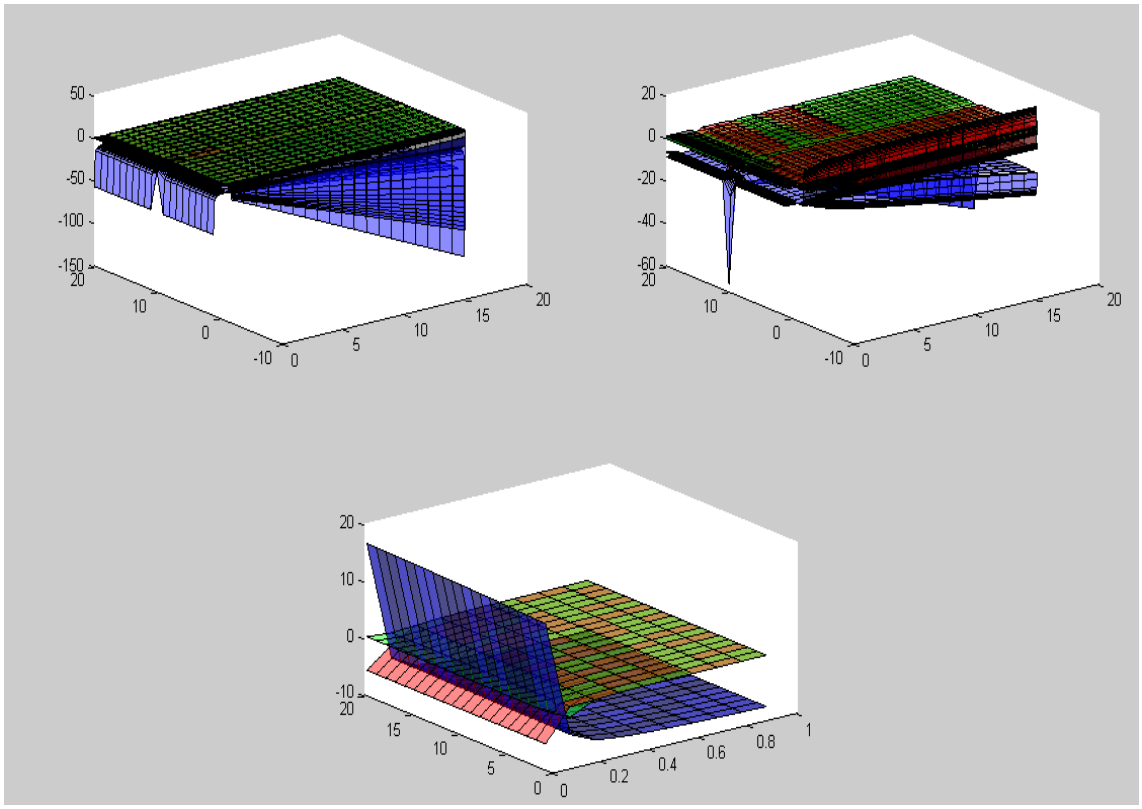
k3 \ k4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
8	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
9	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
10	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
11	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
12	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
13	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
14	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
15	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
16	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
17	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
18	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
19	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
20	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

k_3&k_4’



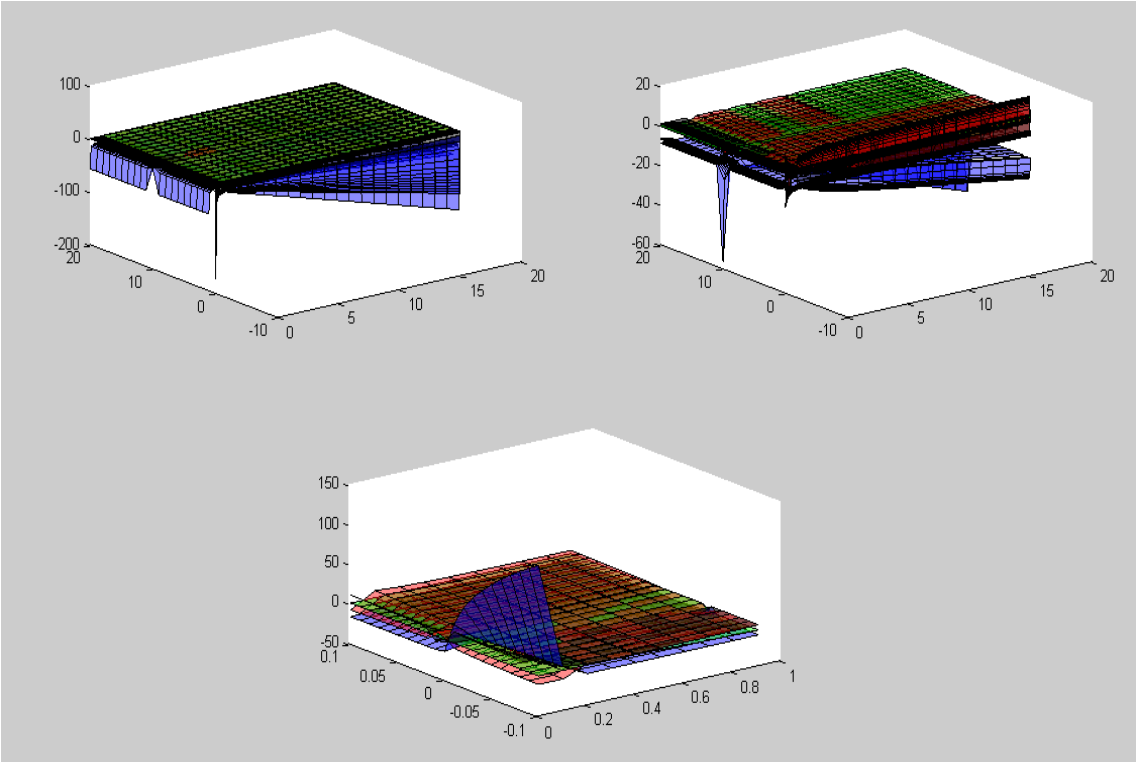
k3 \ k4’	-0.1	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
13	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
14	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
15	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
16	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
17	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
18	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
19	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
20	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0

k_3'&k_4



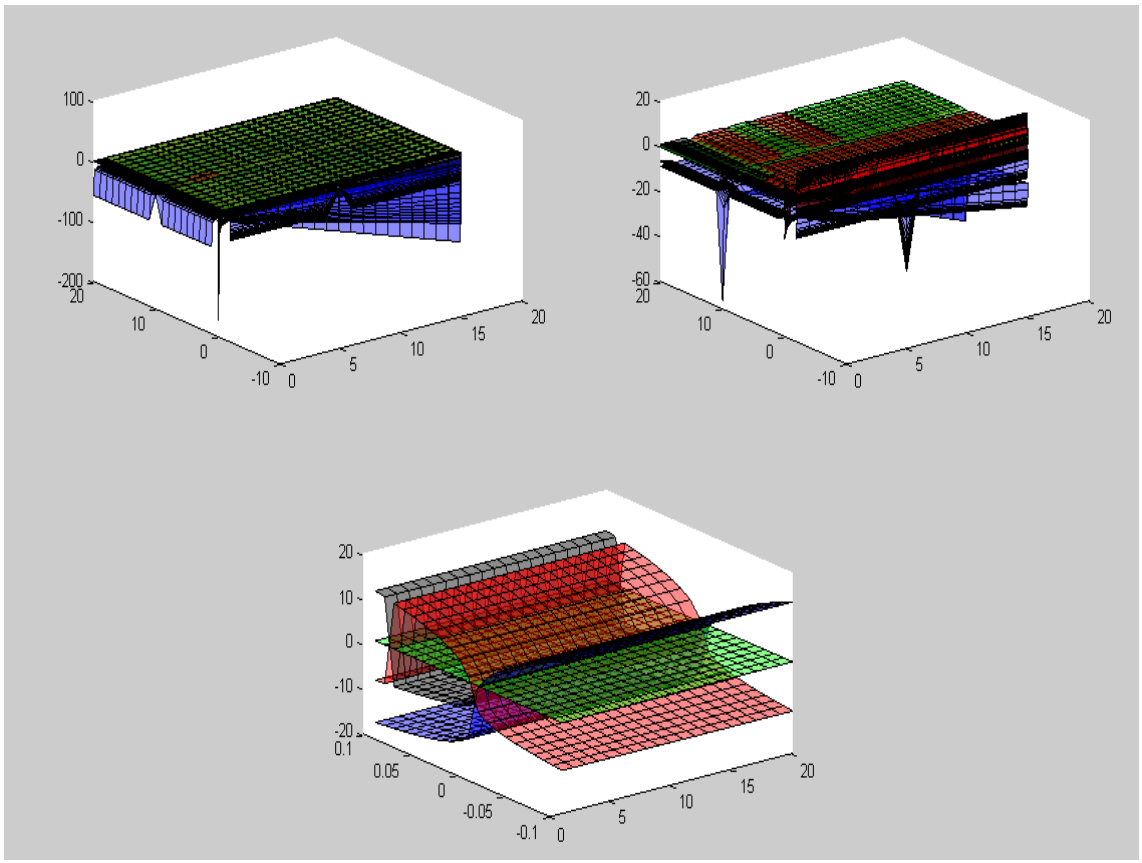
k3' \ k4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.01	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0.3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

k_3'&k_4'



k3' \ k4'	-0.1	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
0.01	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.1	1	1	1	1	1	1	1	1	2	2	0	0	0	0	0	0	0	0	0	0	0
0.2	1	1	1	1	1	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0
0.3	1	1	2	2	2	2	2	2	2	1	0	0	0	0	0	0	0	0	0	0	0
0.4	2	2	2	2	2	2	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.5	2	2	2	2	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.6	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.7	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.8	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0.9	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0

k_4&k_4’



k4 \ k4’	-0.1	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
1	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
13	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
14	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
15	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
16	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
17	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
18	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
19	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0
20	1	1	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0	0	0	0

References

[Acc96] Accili, D., Drago, J., Lee, E. J., Johnson, M. D., Cool, M. H., Salvatore, P., ... & Westphal, H. (1996). Early neonatal death in mice homozygous for a null allele of the insulin receptor gene. *Nature genetics*, *12*(1), 106.

[Acc99] Accili, D., Nakae, J., Kim, J. J., Park, B. C., & Rother, K. I. (1999). Targeted gene mutations define the roles of insulin and IGF-I receptors in mouse embryonic development. *Journal of Pediatric Endocrinology and Metabolism*, *12*(4), 475-486.

[Ani08] Anisimov, V. N., Berstein, L. M., Egormin, P. A., Piskunova, T. S., Popovich, I. G., Zabezhinski, M. A., ... & Semenchenko, A. V. (2008). Metformin slows down aging and extends life span of female SHR mice. *Cell cycle*, *7*(17), 2769-2773.

[Apf04] Apfeld, J., O'Connor, G., McDonagh, T., DiStefano, P. S., & Curtis, R. (2004). The AMP-activated protein kinase AAK-2 links energy levels and insulin-like signals to lifespan in C. elegans. *Genes & development*, *18*(24), 3004-3009.

[Bai91] Bailey, J. E., (1991). Toward a science of metabolic engineering. *Science,* 252 (5013): 1668–1675.

[Bec97] Becker, T., Wullimann, M. F., Becker, C. G., Bernhardt, R. R., & Schachner, M. (1997). Axonal regrowth after spinal cord transection in adult zebrafish. *Journal of Comparative Neurology*, *377*(4), 577-595.

[Ben94] Benevolensky, S. V., Clifton, D., & Fraenkel, D. G. (1994). The effect of increased phosphoglucose isomerase on glucose

metabolism in Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *269*(7), 4878-4882.

[Ben13] Bender, C. M., & Orszag, S. A. (2013). *Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory*. Springer Science & Business Media.

[Ber09] Bersell, K., Arab, S., Haring, B., & Kühn, B. (2009). Neuregulin1/ErbB4 signaling induces cardiomyocyte proliferation and repair of heart injury. *Cell*, *138*(2), 257-270.

[Bin10] Binsl, T. W., Alders, D. J., Heringa, J., Groeneveld, A. B., & Van Beek, J. H. (2010). Computational quantification of metabolic fluxes from a single isotope snapshot: application to an animal biopsy. *Bioinformatics*, *26*(5), 653-660.

[Blu82] Blum, J. J., & Stein, R. B. (1982). On the analysis of metabolic networks. In *Biological regulation and development* (pp. 99-125). Springer, Boston, MA.

[Bol02] Bolster, D. R., Crozier, S. J., Kimball, S. R., & Jefferson, L. S. (2002). AMP-activated protein kinase suppresses protein synthesis in rat skeletal muscle through down-regulated mammalian target of rapamycin (mTOR) signaling. *Journal of Biological Chemistry*, *277*(27), 23977-23980.

[Bru97] Brüning, J. C., Winnay, J., Bonner-Weir, S., Taylor, S. I., Accili, D., & Kahn, C. R. (1997). Development of a novel polygenic model of NIDDM in mice heterozygous for IR and IRS-1 null alleles. *Cell*, *88*(4), 561-572.

[Bru14] Bruder-Nascimento, T., da Silva, M. A., & Tostes, R. C. (2014). The involvement of aldosterone on vascular insulin resistance:

implications in obesity and type 2 diabetes. *Diabetology & metabolic syndrome*, *6*(1), 90.

[Bur01] Burgard, A. P., & Maranas, C. D. (2001). Probing the performance limits of the Escherichia coli metabolic network subject to gene additions or deletions. *Biotechnology and bioengineering*, *74*(5), 364-375.

[Cas86] Casazza, J. P., & Veech, R. L. (1986). The interdependence of glycolytic and pentose cycle intermediates in ad libitum fed rats. *Journal of Biological Chemistry*, *261*(2), 690-698.

[Cha11a] Chan, S. H. J., & Ji, P. (2011). Decomposing flux distributions into elementary flux modes in genome-scale metabolic networks. *Bioinformatics*, *27*(16), 2256-2262.

[Cha11b] Chablais, F., Veit, J., Rainer, G., & Jaźwińska, A. (2011). The zebrafish heart regenerates after cryoinjury-induced myocardial infarction. *BMC developmental biology*, *11*(1), 21.

[Che00] Chen, K. C., Csikasz-Nagy, A., Gyorffy, B., Val, J., Novak, B., & Tyson, J. J. (2000). Kinetic analysis of a molecular model of the budding yeast cell cycle. *Molecular biology of the cell*, *11*(1), 369-391.

[Che05] Chen, W. W., Gao, R. L., & Liu, L. S. (2015). Summary of Chinese Cardiovascular Disease Report, 2014. *Chin Circ J*, *30*(7), 618-622.

[Che08] Chen, H., Xu, C. F., Ma, J., Eliseenkova, A. V., Li, W., Pollock, P. M., ... & Mohammadi, M. (2008). A crystallographic snapshot of tyrosine trans-phosphorylation in action. *Proceedings of the National Academy of Sciences*, *105*(50), 19660-19665.

[Cho13] Choi, V. W. Y., Ng, C. Y. P., Kong, M. K. Y., Cheng, S. H., & Yu, K. N. (2013). Adaptive response to ionising radiation induced by cadmium in zebrafish embryos. *Journal of Radiological Protection*, *33*(1), 101.

[Chr00] Christensen, B., Thykær, J., & Nielsen, J. (2000). Metabolic characterization of high-and low-yielding strains of Penicillium chrysogenum. *Applied microbiology and biotechnology*, *54*(2), 212-217.

[Cle01] Cleeman, J. I., Grundy, S. M., Becker, D., & Clark, L. (2001). Expert panel on detection, evaluation and treatment of high blood cholesterol in adults. Executive summary of the third report of the National Cholesterol Education Program (NCEP) Adult Treatment Panel (ATP III). *Jama*, *285*(19), 2486-2497.

[Con85] Connett, R. J. (1985). In vivo glycolytic equilibria in dog gracilis muscle. *Journal of Biological Chemistry*, *260*(6), 3314-3320.

[Cur95] Curran, M. E., Splawski, I., Timothy, K. W., Vincen, G. M., Green, E. D., & Keating, M. T. (1995). A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell*, *80*(5), 795-803.

[Cur07] Curado, S., Anderson, R. M., Jungblut, B., Mumm, J., Schroeter, E., & Stainier, D. Y. (2007). Conditional targeted cell ablation in zebrafish: a new tool for regeneration studies. *Developmental dynamics: an official publication of the American Association of Anatomists*, *236*(4), 1025-1035.

[Die11] Diep, C. Q., Ma, D., Deo, R. C., Holm, T. M., Naylor, R. W., Arora, N., ... & Zhu, H. (2011). Identification of adult nephron progenitors capable of kidney regeneration in zebrafish. *Nature*, *470*(7332), 95.

[Dil02] Dillin, A., Hsu, A. L., Arantes-Oliveira, N., Lehrer-Graiwer, J., Hsin, H., Fraser, A. G., ... & Kenyon, C. (2002). Rates of behavior and aging specified by mitochondrial function during development. *Science*, *298*(5602), 2398-2401.

[Dry03] Drysch, A., El Massaoudi, M., Mack, C., Takors, R., de Graaf, A. A., & Sahm, H. (2003). Production process monitoring by serial mapping of microbial carbon flux distributions using a novel Sensor Reactor approach: II—13C-labeling-based metabolic flux analysis and l-lysine production. *Metabolic engineering*, *5*(2), 96-107.

[Dur11] Durieux, J., Wolff, S., & Dillin, A. (2011). The cell-non-autonomous nature of electron transport chain-mediated longevity. *Cell*, *144*(1), 79-91.

[Edw00] Edwards, J. S., & Palsson, B. O. (2000). The Escherichia coli MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proceedings of the National Academy of Sciences*, *97*(10), 5528-5533.

[Ekd85] Ekdahl, K. N., & Ekman, P. (1985). Fructose-1, 6-bisphosphatase from rat liver. A comparison of the kinetics of the unphosphorylated enzyme and the enzyme phosphorylated by cyclic AMP-dependent protein kinase. *Journal of Biological Chemistry*, *260*(26), 14173-14179.

[Elo02] Elowitz, M. B., Levine, A. J., Siggia, E. D., & Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, *297*(5584), 1183-1186.

[Env09] Enver, T., Pera, M., Peterson, C., & Andrews, P. W. (2009). Stem cell states, fates, and the rules of attraction. *Cell stem cell*, *4*(5), 387-397.

[Epa06] Epa, V. C., & Ward, C. W. (2006). Model for the complex between the insulin-like growth factor I and its receptor: towards designing antagonists for the IGF-1 receptor. *Protein Engineering, Design and Selection*, *19*(8), 377-384.

[Eve15] Everett, B. M., Brooks, M. M., Vlachos, H. E., Chaitman, B. R., Frye, R. L., & Bhatt, D. L. (2015). Troponin and cardiac events in stable ischemic heart disease and diabetes. *New England Journal of Medicine*, *373*(7), 610-620.

[Far16] Farb, M. G., Karki, S., Park, S. Y., Saggese, S. M., Carmine, B., Hess, D. T., ... & Fuster, J. J. (2016). WNT5A-JNK regulation of vascular insulin resistance in human obesity. *Vascular Medicine*, *21*(6), 489-496.

[Fel86] Fell, D. A., & Small, J. R. (1986). Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochemical Journal*, *238*(3), 781-786.

[Fis04] Fischer, E., Zamboni, N., & Sauer, U. (2004). High-throughput metabolic flux analysis based on gas chromatography–mass spectrometry derived 13C constraints. *Analytical biochemistry*, *325*(2), 308-316.

[Fis05] Fischer, E., & Sauer, U. (2005). Large-scale in vivo flux analysis shows rigidity and suboptimal performance of Bacillus subtilis metabolism. *Nature genetics*, *37*(6), 636.

[Fre71] Freychet, P., Roth, J., & Neville Jr, D. M. (1971). Insulin receptors in the liver: specific binding of [125I] insulin to the plasma membrane and its relation to insulin bioactivity. *Proceedings of the National Academy of Sciences of the United States of America*, *68*(8), 1833.

[Fre12] Freidlin, M. I., & Wentzell, A. D. (2012). Random perturbations. In *Random Perturbations of Dynamical Systems* (pp. 1-28). Springer, Berlin, Heidelberg.

[Fri09] Friedel, A. M., Pike, B. L., & Gasser, S. M. (2009). ATR/Mec1: coordinating fork stability and repair. *Current opinion in cell biology*, *21*(2), 237-244.

[Fuh05] Fuhrer, T., Fischer, E., & Sauer, U. (2005). Experimental identification and quantification of glucose metabolism in seven bacterial species. *Journal of bacteriology*, *187*(5), 1581-1590.

[Gao14] Gao, Y., Wu, F., Zhou, J., Yan, L., Jurczak, M. J., Lee, H. Y., ... & Szendroedi, J. (2014). The H19/let-7 double-negative feedback loop contributes to glucose metabolism in muscle cells. *Nucleic acids research*, *42*(22), 13799-13811.

[Gar85] Gardiner, C. W. (1985). *Handbook of stochastic methods* (Vol. 3, pp. 2-20). Berlin: springer.

[Gar02] Garrity, D. M., Childs, S., & Fishman, M. C. (2002). The heartstrings mutation in zebrafish causes heart/fin Tbx5 deficiency syndrome. *Development*, *129*(19), 4635-4645.

[Gem13] Gemberling, M., Bailey, T. J., Hyde, D. R., & Poss, K. D. (2013). The zebrafish as a model for complex tissue regeneration. *Trends in Genetics*, *29*(11), 611-620.

[Ger02] Gerull, B., Gramlich, M., Atherton, J., McNabb, M., Trombitás, K., Sasse-Klaassen, S., ... & Frenneaux, M. (2002). Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nature genetics*, *30*(2), 201.

[Gil77] Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, *81*(25), 2340-2361.

[Gra99] De Graaf, A. A., Striegel, K., Wittig, R. M., Laufer, B., Schmitz, G., Wiechert, W., ... & Sahm, H. (1999). Metabolic state of Zymomonas mobilis in glucose-, fructose-, and xylose-fed continuous cultures as analysed by 13C-and 31P-NMR spectroscopy. *Archives of microbiology*, *171*(6), 371-385.

[Gre88] Greenfield, N. J., McKenzie, M. A., Adebodun, F., Jordan, F., & Lenard, J. (1988). Metabolism of D-glucose in a wall-less mutant of Neurospora crassa examined by carbon-13 and phosphorus-31 nuclear magnetic resonances: effects of insulin. *Biochemistry*, *27*(23), 8526-8533.

[Gri14] Grivas, J., Haag, M., Johnson, A., Manalo, T., Roell, J., Das, T. L., ... & Lafontant, P. J. (2014). Cardiac repair and regenerative potential in the goldfish (Carassius auratus) heart. *Comparative*

Biochemistry and Physiology Part C: Toxicology & Pharmacology, *163*, 14-23.

[Gom01] Gombert, A. K., dos Santos, M. M., Christensen, B., & Nielsen, J. (2001). Network identification and flux quantification in the central metabolism of Saccharomyces cerevisiae under different conditions of glucose repression. *Journal of bacteriology*, *183*(4), 1441-1451.

[Gon11] González-Rosa, J. M., Martín, V., Peralta, M., Torres, M., & Mercader, N. (2011). Extensive scar formation and regression during heart regeneration after cryoinjury in zebrafish. *Development*, *138*(9), 1663-1674.

[Gon12] González-Rosa, J. M., & Mercader, N. (2012). Cryoinjury as a myocardial infarction model for the study of cardiac regeneration in the zebrafish. *Nature protocols*, *7*(4), 782.

[Gon17] González-Rosa, J. M., Burns, C. E., & Burns, C. G. (2017). Zebrafish heart regeneration: 15 years of discoveries. *Regeneration*, *4*(3), 105-123.

[Gup12] Gupta, V., & Poss, K. D. (2012). Clonally dominant cardiomyocytes direct heart morphogenesis. *Nature*, *484*(7395), 479.

[Hab16] Haberzettl, P., McCracken, J. P., Bhatnagar, A., & Conklin, D. J. (2016). Insulin sensitizers prevent fine particulate matter-induced vascular insulin resistance and changes in endothelial progenitor cell homeostasis. *American Journal of Physiology-Heart and Circulatory Physiology*, *310*(11), H1423-H1438.

[HaB16] Haberzettl, P., O'Toole, T. E., Bhatnagar, A., & Conklin, D. J. (2016). Exposure to fine particulate air pollution causes vascular insulin resistance by inducing pulmonary oxidative stress. *Environmental health perspectives*, *124*(12), 1830-1839.

[Han14] Han, P., Zhou, X. H., Chang, N., Xiao, C. L., Yan, S., Ren, H., ... & Diao, J. P. (2014). Hydrogen peroxide primes heart regeneration with a derepression mechanism. *Cell research*, *24*(9), 1091.

[Har09] Harrison, D. E., Strong, R., Sharp, Z. D., Nelson, J. F., Astle, C. M., Flurkey, K., ... & Pahor, M. (2009). Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *nature*, *460*(7253), 392.

[Hein15] Hein, S. J., Lehmann, L. H., Kossack, M., Juergensen, L., Fuchs, D., Katus, H. A., & Hassel, D. (2015). Advanced echocardiography in adult zebrafish reveals delayed recovery of heart function after myocardial cryoinjury. *PLoS One*, *10*(4), e0122665.

[Hin15] Hindupur, S. K., González, A., & Hall, M. N. (2015). The opposing actions of target of rapamycin and AMP-activated protein kinase in cell growth control. *Cold Spring Harbor perspectives in biology*, *7*(8), a019141.

[Hub94] Hubbard, S. R., Wei, L., & Hendrickson, W. A. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature*, *372*(6508), 746.

[Huo14] Huo, Y., & Ji, P. (2014). Continuous-Time Markov Chain–Based Flux Analysis in Metabolism. *Journal of Computational Biology*, *21*(9), 691-698.

[Imp12] Imperatore, G., Boyle, J. P., Thompson, T. J., Case, D., Dabelea, D., Hamman, R. F., ... & Rodriguez, B. L. (2012). Projections of type 1 and type 2 diabetes burden in the US population aged< 20 years through 2050: dynamic modeling of incidence, mortality, and population growth. *Diabetes care*, *35*(12), 2515-2520.

[Ito14] Ito, K., Morioka, M., Kimura, S., Tasaki, M., Inohaya, K., & Kudo, A. (2014). Differential reparative phenotypes between zebrafish and medaka after cardiac injury. *Developmental Dynamics*, *243*(9), 1106-1115.

[Jam02] James, A., Brindley, J., & McIntosh, A. C. (2002). Stability of multiple steady states of catalytic combustion. *Combustion and flame*, *130*(1-2), 137-146.

[Jay08] Jayawardhana, B., Kell, D. B., & Rattray, M. (2008). Bayesian inference of the sites of perturbations in metabolic pathways via Markov chain Monte Carlo. *Bioinformatics*, *24*(9), 1191-1197.

[Jia04] Jia, K., Chen, D., & Riddle, D. L. (2004). The TOR pathway interacts with the insulin signaling pathway to regulate C. elegans larval development, metabolism and life span. *Development*, *131*(16), 3897-3906.

[Jia18] Jia, G., Whaley-Connell, A., & Sowers, J. R. (2018). Diabetic cardiomyopathy: a hyperglycaemia-and insulin-resistance-induced heart disease. *Diabetologia*, *61*(1), 21-28.

[Joh03] John, G. T., Klimant, I., Wittmann, C., & Heinzle, E. (2003). Integrated optical sensing of dissolved oxygen in microtiter plates: a novel tool for microbial cultivation. *Biotechnology and bioengineering*, *81*(7), 829-836.

[Jop10] Jopling, C., Sleep, E., Raya, M., Martí, M., Raya, A., & Belmonte, J. C. I. (2010). Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature*, *464*(7288), 606.

[Kad06] Kadirkamanathan, V., Yang, J., Billings, S. A., & Wright, P. C. (2006). Markov Chain Monte Carlo Algorithm based metabolic flux distribution analysis on Corynebacterium glutamicum. *Bioinformatics*, *22*(21), 2681-2687.

[Kae05] Kaeberlein, M., Powers, R. W., Steffen, K. K., Westman, E. A., Hu, D., Dang, N., ... & Kennedy, B. K. (2005). Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients. *Science*, *310*(5751), 1193-1196.

[Kan16] Kang, J., Hu, J., Karra, R., Dickson, A. L., Tornini, V. A., Nachtrab, G., ... & Poss, K. D. (2016). Modulation of tissue repair by regeneration enhancer elements. *Nature*, *532*(7598), 201.

[Kee98] Keener, J. P., & Sneyd, J. (1998). *Mathematical physiology*(Vol. 1). New York: Springer.

[Kee09] Keener, J. P., & Sneyd, J. (2009). *Mathematical physiology 1: Cellular physiology* (Vol. 2). Springer.

[Ken93] Kenyon, C., Chang, J., Gensch, E., Rudner, A., & Tabtiang, R. (1993). A C. elegans mutant that lives twice as long as wild type. *Nature*, *366*(6454), 461.

[Ken05] Kenyon, C. (2005). The plasticity of aging: insights from long-lived mutants. *Cell*, *120*(4), 449-460.

[Kid00] Kido, Y., Burks, D. J., Withers, D., Bruning, J. C., Kahn, C. R., White, M. F., & Accili, D. (2000). Tissue-specific insulin resistance in mice with mutations in the insulin receptor, IRS-1, and IRS-2. *The Journal of clinical investigation*, *105*(2), 199-205.

[Kik10] Kikuchi, K., Holdway, J. E., Werdich, A. A., Anderson, R. M., Fang, Y., Egnaczyk, G. F., ... & Poss, K. D. (2010). Primary contribution to zebrafish heart regeneration by gata4+ cardiomyocytes. *Nature*, *464*(7288), 601.

[Kim97] Kimura, K. D., Tissenbaum, H. A., Liu, Y., & Ruvkun, G. (1997). daf-2, an insulin receptor-like gene that regulates longevity and diapause in Caenorhabditis elegans. *Science*, *277*(5328), 942-946.

[Kla83] Klass, M. R. (1983). A method for the isolation of longevity mutants in the nematode Caenorhabditis elegans and initial results. *Mechanisms of ageing and development*, *22*(3-4), 279-286.

[Kle06] Kleijn, R. J., van Winden, W. A., Ras, C., van Gulik, W. M., Schipper, D., & Heijnen, J. J. (2006). 13C-labeled gluconate tracing as a direct and accurate method for determining the

pentose phosphate pathway split ratio in Penicillium chrysogenum. *Appl. Environ. Microbiol.*, *72*(7), 4743-4754.

[Koj04] Kojima, T., Kamei, H., Aizu, T., Arai, Y., Takayama, M., Nakazawa, S., ... & Sakaki, Y. (2004). Association analysis between longevity in the Japanese population and polymorphic variants of genes involved in insulin and insulin-like growth factor 1 signaling pathways. *Experimental gerontology*, *39*(11-12), 1595-1598.

[Kro11] Kroehne, V., Freudenreich, D., Hans, S., Kaslin, J., & Brand, M. (2011). Regeneration of the adult zebrafish brain from neurogenic radial glia-type progenitors. *Development*, *138*(22), 4831-4841.

[Kul99] Kulkarni, R. N., Brüning, J. C., Winnay, J. N., Postic, C., Magnuson, M. A., & Kahn, C. R. (1999). Tissue-specific knockout of the insulin receptor in pancreatic β cells creates an insulin secretory defect similar to that in type 2 diabetes. *Cell*, *96*(3), 329-339.

[Laf12] Lafontant, P. J., Burns, A. R., Grivas, J. A., Lesch, M. A., Lala, T. D., Reuter, S. P., ... & Frounfelter, T. D. (2012). The giant danio (D. aequipinnatus) as a model of cardiac remodeling and regeneration. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, *295*(2), 234-248.

[Lan03] Langheinrich, U., Vacun, G., & Wagner, T. (2003). Zebrafish embryos express an orthologue of HERG and are sensitive toward a range of QT-prolonging drugs inducing severe arrhythmia. *Toxicology and applied pharmacology*, *193*(3), 370-382.

[Lep06] Lepilina, A., Coon, A. N., Kikuchi, K., Holdway, J. E., Roberts, R. W., Burns, C. G., & Poss, K. D. (2006). A dynamic epicardial injury response supports progenitor cell activity during zebrafish heart regeneration. *Cell*, *127*(3), 607-619.

[Ler15] Leri, A., Rota, M., Pasqualini, F. S., Goichberg, P., & Anversa, P. (2015). Origin of cardiomyocytes in the adult heart. *Circulation research*, *116*(1), 150-166.

[Lev50] Levine, R., Goldstein, M. S., Huddlestun, B., & Klein, S. P. (1950). Action of insulin on the 'permeability'of cells to free hexoses, as studied by its effect on the distribution of galactose. *American Journal of Physiology-Legacy Content*, *163*(1), 70-76.

[Lie06] Lien, C. L., Schebesta, M., Makino, S., Weber, G. J., & Keating, M. T. (2006). Gene expression analysis of zebrafish heart regeneration. *PLoS biology*, *4*(8), e260.

[LiG11] Li, G. W., & Xie, X. S. (2011). Central dogma at the single-molecule level in living cells. *Nature*, *475*(7356), 308.

[Lin97] Lin, K., Dorman, J. B., Rodan, A., & Kenyon, C. (1997). daf-16: An HNF-3/forkhead family member that can function to double the life-span of Caenorhabditis elegans. *Science*, *278*(5341), 1319-1322.

[LiQ97] Li, Q. Y., Newbury-Ecob, R. A., Terrett, J. A., Wilson, D. I., Curtis, A. R., Yi, C. H., ... & Bonnet, D. (1997). Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nature genetics*, *15*(1), 21.

[Liu05] Liu, X., Jiang, N., Hughes, B., Bigras, E., Shoubridge, E., & Hekimi, S. (2005). Evolutionary conservation of the clk-1-dependent mechanism of longevity: loss of mclk1 increases cellular fitness and lifespan in mice. *Genes & development*, *19*(20), 2424-2434.

[Liu12] Liu, J., & Stainier, D. Y. (2012). Zebrafish in the study of early cardiac development. *Circulation research*, *110*(6), 870-874.

[Lou06] Lou, M., Garrett, T. P., McKern, N. M., Hoyne, P. A., Epa, V. C., Bentley, J. D., ... & Ward, C. W. (2006). The first three domains of the insulin receptor differ structurally from the insulin-like growth factor 1 receptor in the regions governing ligand specificity. *Proceedings of the National Academy of Sciences*, *103*(33), 12429-12434.

[Luo07] Luo, B., Groenke, K., Takors, R., Wandrey, C., & Oldiges, M. (2007). Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography–mass spectrometry. *Journal of chromatography A*, *1147*(2), 153-164.

[Mae14] Maejima, Y., Chen, Y., Isobe, M., Gustafsson, Å. B., Kitsis, R. N., & Sadoshima, J. (2014). Recent progress in research on molecular mechanisms of autophagy in the heart. *American Journal of Physiology-Heart and Circulatory Physiology*, *308*(4), H259-H268.

[Mah02] Mahadevan, R., Edwards, J. S., & Doyle III, F. J. (2002). Dynamic flux balance analysis of diauxic growth in Escherichia coli. *Biophysical journal*, *83*(3), 1331-1340.

[Mal88] Malloy, C. R., Sherry, A. D., & Jeffrey, F. M. (1988). Evaluation of carbon flux and substrate selection through alternate pathways involving the citric acid cycle of the heart by 13C NMR spectroscopy. *Journal of Biological Chemistry*, *263*(15), 6964-6971.

[Mar55] Maruyama, G. (1955). Continuous Markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo*, *4*(1), 48-90.

[Mar85] Marshall, S. (1985). Kinetics of insulin receptor internalization and recycling in adipocytes. Shunting of receptors to a degradative pathway by inhibitors of recycling. *Journal of Biological Chemistry*, *260*(7), 4136-4144.

[Mar96] Marx, A., Erdmann, P., Senn, M., Körner, S., Jungo, T., Petretta, M., ... & Zehnder, M. (1996). Synthesis of 4′-C-Acylated Thymidines. *Helvetica chimica acta*, *79*(7), 1980-1994.

[McK06] McKern, N. M., Lawrence, M. C., Streltsov, V. A., Lou, M. Z., Adams, T. E., Lovrecz, G. O., ... & Hoyne, P. A. (2006). Structure of the insulin receptor ectodomain reveals a folded-over conformation. *Nature*, *443*(7108), 218.

[Mez12] Meza, E., Becker, J., Bolivar, F., Gosset, G., & Wittmann, C. (2012). Consequences of phosphoenolpyruvate: sugar phosphotranferase system and pyruvate kinase isozymes inactivation in central carbon metabolism flux distribution in Escherichia coli. *Microbial cell factories*, *11*(1), 127.

[Mic01] Miclet, E., Stoven, V., Michels, P. A., Opperdoes, F. R., Lallemand, J. Y., & Duffieux, F. (2001). NMR spectroscopic

analysis of the first two steps of the pentose-phosphate pathway elucidates the role of 6-phosphogluconolactonase. *Journal of Biological Chemistry*, *276*(37), 34840-34846.

[Mic17] Micha, R., Peñalvo, J. L., Cudhea, F., Imamura, F., Rehm, C. D., & Mozaffarian, D. (2017). Association between dietary factors and mortality from heart disease, stroke, and type 2 diabetes in the United States. *Jama*, *317*(9), 912-924.

[Mil79] Milstein, J., & Bremermann, H. J. (1979). Parameter identification of the Calvin photosynthesis cycle. *Journal of Mathematical Biology*, *7*(2), 99-116.

[Mos10] Moscou, S. (2010). Getting the word out: advocacy, social marketing, and policy development and enforcement. *Public Health Nursing*, 285.

[Mun07] Muniyappa, R., Montagnani, M., Koh, K. K., & Quon, M. J. (2007). Cardiovascular actions of insulin. *Endocrine reviews*, *28*(5), 463-491.

[Nef03] Nef, S., Verma-Kurvari, S., Merenmies, J., Vassalli, J. D., Efstratiadis, A., Accili, D., & Parada, L. F. (2003). Testis determination requires insulin receptor family function in mice. *Nature*, *426*(6964), 291.

[Pap84] ET Papoutsakis. (1984). Equations and calculations for fermentations of butyric acid bacteria. *Biotechnol Bioeng,* 26(2):174-87

[Par00] Park, S. Y., Kim, H. K., Yoo, S. K., Oh, T. K., & Lee, J. K. (2000). Characterization of glk, a gene coding for glucose kinase

of Corynebacterium glutamicum. *FEMS Microbiology letters*, *188*(2), 209-215.

[Par13] Parente, V., Balasso, S., Pompilio, G., Verduci, L., Colombo, G. I., Milano, G., ... & Capogrossi, M. C. (2013). Hypoxia/reoxygenation cardiac injury and regeneration in zebrafish adult heart. *PloS one*, *8*(1), e53748.

[Paw09] Pawlikowska, L., Hu, D., Huntsman, S., Sung, A., Chu, C., Chen, J., ... & Psaty, B. M. (2009). Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. *Aging cell*, *8*(4), 460-472.

[Pfe15] Pfefferli, C., & Jaźwińska, A. (2015). The art of fin regeneration in zebrafish. *Regeneration*, *2*(2), 72-83.

[Pfe17] Pfefferli, C., & Jaźwińska, A. (2017). The careg element reveals a common regulation of regeneration in the zebrafish myocardium and fin. *Nature communications*, *8*, 15151.

[Pet00] Petersen, S., de Graaf, A. A., Eggeling, L., Möllney, M., Wiechert, W., & Sahm, H. (2000). In vivo quantification of parallel and bidirectional fluxes in the anaplerosis of Corynebacterium glutamicum. *Journal of Biological Chemistry*, *275*(46), 35932-35941.

[Por11] Porrello, E. R., Mahmoud, A. I., Simpson, E., Hill, J. A., Richardson, J. A., Olson, E. N., & Sadek, H. A. (2011). Transient regenerative potential of the neonatal mouse heart. *Science*, *331*(6020), 1078-1080.

[Pos02] Poss, K. D., Wilson, L. G., & Keating, M. T. (2002). Heart regeneration in zebrafish. *Science*, *298*(5601), 2188-2190.

[Pos03] Poss, K. D., Nechiporuk, A., Johnson, S. L., & Keating, M. T. (2003, July). Heart and fin regeneration in zebrafish. In *DEVELOPMENTAL BIOLOGY* (Vol. 259, No. 2, pp. 525-525).

[Pra97] Pramanik, J., & Keasling, J. D. (1997). Stoichiometric model of Escherichia coli metabolism: incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnology and bioengineering*, *56*(4), 398-421.

[Ray03] Raya, Á., Koth, C. M., Büscher, D., Kawakami, Y., Itoh, T., Raya, R. M., ... & Izpisúa-Belmonte, J. C. (2003). Activation of Notch signaling pathway precedes heart regeneration in zebrafish. *Proceedings of the National Academy of Sciences*, *100*(suppl 1), 11889-11895.

[Rem05] Remacle-Bonnet, M., Garrouste, F., Baillat, G., Andre, F., Marvaldi, J., & Pommier, G. (2005). Membrane rafts segregate pro-from anti-apoptotic insulin-like growth factor-I receptor signaling in colon carcinoma cells stimulated by members of the tumor necrosis factor superfamily. *The American journal of pathology*, *167*(3), 761-773.

[Ren08] Rentería, M. E., Gandhi, N. S., Vinuesa, P., Helmerhorst, E., & Mancera, R. L. (2008). A comparative structural bioinformatics analysis of the insulin receptor family ectodomain based on phylogenetic information. *PloS one*, *3*(11), e3667.

[Ric17] Richards, D. J., Coyle, R. C., Tan, Y., Jia, J., Wong, K., Toomer, K., ... & Mei, Y. (2017). Inspiration from heart development:

Biomimetic development of functional human cardiac organoids. *Biomaterials*, *142*, 112-123.

[Rod16] Rodius, S., Androsova, G., Götz, L., Liechti, R., Crespo, I., Merz, S., ... & Muller, A. (2016). Analysis of the dynamic co-expression network of heart regeneration in the zebrafish. *Scientific reports*, *6*, 26822.

[Sar97] Saraste, A., Pulkki, K., Kallajoki, M., Henriksen, K., Parvinen, M., & Voipio-Pulkki, L. M. (1997). Apoptosis in human acute myocardial infarction. *Circulation*, *95*(2), 320-323.

[Sau99] Sauer, U., & Bailey, J. E. (1999). Estimation of P-to-O ratio in Bacillus subtilis and its influence on maximum riboflavin yield. *Biotechnology and bioengineering*, *64*(6), 750-754.

[Sav92] Savinell, J. M., & Palsson, B. O. (1992). Optimal selection of metabolic fluxes for in vivo measurement. II. Application to Escherichia coli and hybridoma cell metabolism. *Journal of theoretical biology*, *155*(2), 215-242.

[Sav92] Savinell, J. M., & Palsson, B. O. (1992). Optimal selection of metabolic fluxes for in vivo measurement. I. Development of mathematical methods. *Journal of theoretical biology*, *155*(2), 201-214.

[Sel04] Selivanov, V. A., Puigjaner, J., Sillero, A., Centelles, J. J., Ramos-Montoya, A., Lee, P. W. N., & Cascante, M. (2004). An optimized algorithm for flux estimation from isotopomer distribution in glucose metabolites. *Bioinformatics*, *20*(18), 3387-3397.

[Sch99] Schmidt, K., Nielsen, J., & Villadsen, J. (1999). Quantitative analysis of metabolic fluxes in Escherichia coli, using two-dimensional NMR spectroscopy and complete isotopomer models. *Journal of Biotechnology*, *71*(1-3), 175-189.

[Sch04] Schulze, P. C., & Lee, R. T. (2004). Macrophage-mediated cardiac fibrosis. *Circulation research*, *95*(6), 552-553.

[Sch11] Schnabel, K., Wu, C. C., Kurth, T., & Weidinger, G. (2011). Regeneration of cryoinjury induced necrotic heart lesions in zebrafish is associated with epicardial activation and cardiomyocyte proliferation. *PloS one*, *6*(4), e18503.

[Sch12] Schellenberger, J., Zielinski, D. C., Choi, W., Madireddi, S., Portnoy, V., Scott, D. A., ... & Osterman, A. L. (2012). Predicting outcomes of steady-state ^{13}C isotope tracing experiments using Monte Carlo sampling. *BMC systems biology*, *6*(1), 9.

[Sch15] Schulz, A. M., & Haynes, C. M. (2015). UPRmt-mediated cytoprotection and organismal aging. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1847*(11), 1448-1456.

[Sel04] Selivanov, V. A., Puigjaner, J., Sillero, A., Centelles, J. J., Ramos-Montoya, A., Lee, P. W. N., & Cascante, M. (2004). An optimized algorithm for flux estimation from isotopomer distribution in glucose metabolites. *Bioinformatics*, *20*(18), 3387-3397.

[Sla06] Slaaby, R., Schäffer, L., Lautrup-Larsen, I., Andersen, A. S., Shaw, A. C., Mathiasen, I. S., & Brandt, J. (2006). Hybrid receptors formed by insulin receptor (IR) and insulin-like

growth factor I receptor (IGF-IR) have low insulin and high IGF-1 affinity irrespective of the IR splice variant. *Journal of Biological Chemistry*, *281*(36), 25869-25874.

[Sle10] Sleep, E., Boue, S., Jopling, C., Raya, M., Raya, A., & Belmonte, J. C. I. (2010). Transcriptomics approach to investigate zebrafish heart regeneration. *Journal of cardiovascular medicine*, *11*(5), 369-380.

[Smy06] Smyth, S., & Heron, A. (2006). Diabetes and obesity: the twin epidemics. *Nature medicine*, *12*(1), 75.

[Soo98] Soonpaa, M. H., & Field, L. J. (1998). Survey of studies examining mammalian cardiomyocyte DNA synthesis. *Circulation research*, *83*(1), 15-26.

[Sor90] Sorrentino, B., Ney, P., Bodine, D., & Nienhius, A. W. (1990). A 46 base pair enhancer sequence within the locus activating region is required for induced expression of the gamma-globin gene during erythroid differentiation. *Nucleic Acids Research*, *18*(9), 2721-2731.

[Ste98] Stephanopoulos, G., Aristidou, A. A., & Nielsen, J. (1998). *Metabolic engineering: principles and methodologies*. Elsevier.

[Sun16] Sun, N., Youle, R. J., & Finkel, T. (2016). The mitochondrial basis of aging. *Molecular cell*, *61*(5), 654-666.

[Tan17] Tan, C., Wang, A., Liu, C., Li, Y., Shi, Y., & Zhou, M. S. (2017). Puerarin improves vascular insulin resistance and cardiovascular remodeling in salt-sensitive hypertension. *The American journal of Chinese medicine*, *45*(06), 1169-1184.

[Tav93] Tavaré, J. M., & Siddle, K. (1993). Mutational analysis of insulin receptor function: consensus and controversy. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1178*(1), 21-39.

[Tay98] Taylor, S. I., & Arioglu, E. (1998). Syndromes associated with insulin resistance and acanthosis nigricans. *Journal of basic and clinical physiology and pharmacology*, *9*(2-4), 419-439.

[Tfa09] Tfayli, H., & Arslanian, S. (2009). Pathophysiology of type 2 diabetes mellitus in youth: the evolving chameleon. *Arquivos Brasileiros de Endocrinologia & Metabologia*, *53*(2), 165-174.

[Thi08] Thiel, C. T., Knebel, B., Knerr, I., Sticht, H., Müller-Wieland, D., Zenker, M., ... & Rauch, A. (2008). Two novel mutations in the insulin binding subunit of the insulin receptor gene without insulin binding impairment in a patient with Rabson–Mendenhall syndrome. *Molecular genetics and metabolism*, *94*(3), 356-362.

[Tia15] Tian, S., Liu, Q., Gnatovskiy, L., Ma, P. X., & Wang, Z. (2015). Heart regeneration with embryonic cardiac progenitor cells and cardiac tissue engineering. *Journal of stem cell and transplantation biology*, *1*(1).

[Tou13] Tousoulis, D., Papageorgiou, N., Androulakis, E., Siasos, G., Latsios, G., Tentolouris, K., & Stefanadis, C. (2013). Diabetes mellitus-associated vascular impairment: novel circulating biomarkers and therapeutic approaches. *Journal of the American College of Cardiology*, *62*(8), 667-676.

[Tri04] Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J. N., Rovio, A. T., Bruder, C. E., ... & Törnell, J. (2004). Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature*, *429*(6990), 417.

[Val93] Vallino, J. J., & Stephanopoulos, G. (2000). Metabolic flux distributions in Corynebacterium glutamicum during growth and lysine overproduction. *Biotechnology and bioengineering*, *67*(6), 872-885.

[Var93] Varma, A., Boesch, B. W., & Palsson, B. O. (1993). Biochemical production capabilities of Escherichia coli. *Biotechnology and bioengineering*, *42*(1), 59-73.

[Vee69] Veech, R. L., Raijman, L., Dalziel, K., & Krebs, H. A. (1969). Disequilibrium in the triose phosphate isomerase system in rat liver. *Biochemical Journal*, *115*(4), 837-842.

[Vih00] Vihtelic, T. S., & Hyde, D. R. (2000). Light-induced rod and cone cell death and regeneration in the adult albino zebrafish (Danio rerio) retina. *Journal of neurobiology*, *44*(3), 289-307.

[Wan11] Wang, J., Panáková, D., Kikuchi, K., Holdway, J. E., Gemberling, M., Burris, J. S., ... & Werdich, A. A. (2011). The regenerative capacity of zebrafish reverses cardiac failure caused by genetic cardiomyocyte depletion. *Development*, *138*(16), 3421-3430.

[War04] Ward, C. W., & Garrett, T. P. (2004). Structural relationships between the insulin receptor and epidermal growth factor receptor families and other proteins. *Current opinion in drug discovery & development*, *7*(5), 630-638.

[Wat53] Watson, J. D. and Crick, F. H. C. (1953) The structure of DNA. *Cold Spring Harbor symposia on Quantitative Biology*, 18: 123-131.

[Wen09] Wenz, T., Rossi, S. G., Rotundo, R. L., Spiegelman, B. M., & Moraes, C. T. (2009). Increased muscle PGC-1α expression protects from sarcopenia and metabolic disease during aging. *Proceedings of the National Academy of Sciences*, *106*(48), 20405-20410.

[Whi03] White, M. F. (2003). Insulin signaling in health and disease. *Science*, *302*(5651), 1710-1711.

[Wie97] Wiechert, W., Siefke, C., de Graaf, A. A., & Marx, A. (1997). Bidirectional reaction steps in metabolic networks: II. Flux estimation and statistical analysis. *Biotechnology and bioengineering*, *55*(1), 118-135.

[Win02] van Winden, W. A., Heijnen, J. J., & Verheijen, P. J. (2002). Cumulative bondomers: a new concept in flux analysis from 2D [13C, 1H] COSY NMR data. *Biotechnology and bioengineering*, *80*(7), 731-745.

[WuC16] Wu, C. C., Kruse, F., Vasudevarao, M. D., Junker, J. P., Zebrowski, D. C., Fischer, K., ... & Van Oudenaarden, A. (2016). Spatially resolved genome-wide transcriptional profiling identifies BMP signaling as essential regulator of zebrafish cardiomyocyte regeneration. *Developmental cell*, *36*(1), 36-49.

[WuJ96] Wu, J., Garami, M., Cheng, T., & Gardner, D. G. (1996). 1, 25 (OH) 2 vitamin D3, and retinoic acid antagonize endothelin-

stimulated hypertrophy of neonatal rat cardiac myocytes. *The Journal of clinical investigation*, *97*(7), 1577-1588.

[XuX02] Xu, X., Meiler, S. E., Zhong, T. P., Mohideen, M., Crossley, D. A., Burggren, W. W., & Fishman, M. C. (2002). Cardiomyopathy in zebrafish due to mutation in an alternatively spliced exon of titin. *Nature genetics*, *30*(2), 205.

[XuY13] Xu, Y., Wang, L., He, J., Bi, Y., Li, M., Wang, T., ... & Xu, M. (2013). Prevalence and control of diabetes in Chinese adults. *Jama*, *310*(9), 948-959.

[XuZ13] Xu, Z., Sun, X., & Sun, J. (2013). Construction and analysis of the model of energy metabolism in E. coli. *PloS one*, *8*(1), e55137.

[YuQ10] Yu, Q., Gao, F., & Ma, X. L. (2010). Insulin says NO to cardiovascular disease. *Cardiovascular research*, *89*(3), 516-524.

[Zha08] Zhao, Z., Kuijvenhoven, K., Ras, C., van Gulik, W. M., Heijnen, J. J., Verheijen, P. J., & van Winden, W. A. (2008). Isotopic non-stationary 13C gluconate tracer method for accurate determination of the pentose phosphate pathway split-ratio in Penicillium chrysogenum. *Metabolic Engineering*, *10*(3-4), 178-186.

[Zha13a] Zhang, Y. (2013). A dynamic Bayesian Markov model for phasing and characterizing haplotypes in next-generation sequencing. *Bioinformatics*, *29*(7), 878-885.

[Zha13b] Zhang, R., Han, P., Yang, H., Ouyang, K., Lee, D., Lin, Y. F., ... & Yelon, D. (2013). In vivo cardiac reprogramming contributes to zebrafish heart regeneration. *Nature*, *498*(7455), 497.

[Zhe12] Zhen, Y. S., Wu, Q., Xiao, C. L., Chang, N. N., Wang, X., Lei, L., ... & Xiong, J. W. (2012). Overlapping cardiac programs in heart development and regeneration. *Journal of genetics and genomics*, *39*(9), 443-449.

[Zic05] Zick, Y. (2005). Ser/Thr phosphorylation of IRS proteins: a molecular basis for insulin resistance. *Sci. STKE*, *2005*(268), pe4-pe4.

[Zup95] Zupke, C., & Stephanopoulos, G. (1995). Intracellular flux analysis in hybridomas using mass balances and in vitro 13C NMR. *Biotechnology and bioengineering*, *45*(4), 292-303.