

# IOWA STATE UNIVERSITY

## Digital Repository

---

Graduate Theses and Dissertations

Iowa State University Capstones, Theses and  
Dissertations

---

2015

## Mathematical and statistical models in evolutionary game theory

Min Wang  
*Iowa State University*

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>



Part of the [Applied Mathematics Commons](#), and the [Statistics and Probability Commons](#)

---

### Recommended Citation

Wang, Min, "Mathematical and statistical models in evolutionary game theory" (2015). *Graduate Theses and Dissertations*. 14449.  
<https://lib.dr.iastate.edu/etd/14449>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

**Mathematical and statistical models in evolutionary game theory**

by

**Min Wang**

A dissertation submitted to the graduate faculty  
in partial fulfillment of the requirements for the degree of  
**DOCTOR OF PHILOSOPHY**

Co-majors: Applied Mathematics  
Statistics

Program of Study Committee:  
Zhijun Wu, Co-major Professor  
Karin Dorman, Co-major Professor  
Stephen Willson  
Daniel Nordman  
Alexander Roitershtein

Iowa State University  
Ames, Iowa  
2015  
Copyright © Min Wang, 2015. All rights reserved.

**DEDICATION**

*This thesis is dedicated to my family. To my parents, who have raised me to the person I am today. To my younger sister, who has kept the family full of happiness when I am far away.*

## TABLE OF CONTENTS

|   |             |
|---|-------------|
| <b>LIST OF TABLES . . . . .</b>                       | <b>vii</b>  |
| <b>LIST OF FIGURES . . . . .</b>                      | <b>viii</b> |
| <b>ACKNOWLEDGEMENTS . . . . .</b>                     | <b>xiv</b>  |
| <b>ABSTRACT . . . . .</b>                             | <b>xv</b>   |
| <b>CHAPTER 1. INTRODUCTION . . . . .</b>              | <b>1</b>    |
| 1.1 Evolutionary Game Theory . . . . .                | 1           |
| 1.1.1 Evolutionary Theory . . . . .                   | 1           |
| 1.1.2 Game Theory . . . . .                           | 2           |
| 1.2 Outline of the Thesis . . . . .                   | 2           |
| <b>CHAPTER 2. 2D YEAST COOPERATION GAME . . . . .</b> | <b>4</b>    |
| 2.1 Introduction . . . . .                            | 4           |
| 2.2 Yeast System and the Nonlinear Payoffs . . . . .  | 7           |
| 2.3 Linear Games and Simulations . . . . .            | 9           |
| 2.4 Nonlinear Games and 2D Simulations . . . . .      | 13          |
| 2.4.1 Deterministic Rule . . . . .                    | 17          |
| 2.4.2 Stochastic Rule . . . . .                       | 19          |
| 2.5 General Mathematical Theory . . . . .             | 21          |
| 2.5.1 Defectors Invading Cooperators . . . . .        | 22          |
| 2.5.2 Cooperators Invading Defectors . . . . .        | 24          |
| 2.5.3 Generalization . . . . .                        | 27          |
| 2.6 Conclusion and Discussion . . . . .               | 28          |

|  |    |
|--|----|
| <b>CHAPTER 3. ESS IN NONLINEAR GAME AND THE ASYMPTOTIC BEHAVIOR OF<br/>SOLUTION TO ADAPTED REPLICATOR DIFFUSION EQUATION</b> . . . . . | 32 |
| 3.1 Introduction . . . . .   | 32 |
| 3.2 ESS in Nonlinear Games . . . . .   | 35 |
| 3.3 Long Time Behavior of Solution to the Adapted Replicator Diffusion Equation . . . . .  | 39 |
| 3.4 Examples and Numerical Simulation Results . . . . .  | 45 |
| 3.4.1 Constructed Example . . . . .  | 48 |
| 3.4.2 Yeast Example . . . . .  | 50 |
| 3.5 Discussion and Conclusions . . . . .   | 55 |
| <b>CHAPTER 4. PARAMETER ESTIMATION IN DYNAMICAL METAGENOMIC MODEL</b> . . . . .  | 57 |
| 4.1 Introduction . . . . .   | 57 |
| 4.2 Motivation . . . . .   | 60 |
| 4.3 Model . . . . .  | 62 |
| 4.3.1 Identifiability Analysis . . . . .   | 62 |
| 4.3.2 Dispersion Estimator . . . . .   | 64 |
| 4.3.3 NB Mixture Model . . . . .   | 65 |
| 4.3.4 EM Algorithm . . . . .   | 68 |
| 4.3.5 Initialization . . . . .   | 69 |
| 4.4 Simulation Study . . . . .   | 70 |
| 4.4.1 Data Simulation . . . . .  | 70 |
| 4.4.2 Results . . . . .  | 72 |
| 4.5 Discussion and Future Work . . . . .   | 75 |
| <b>CHAPTER 5. MARKOV RANDOM FIELD MODEL FOR ANALYSIS OF GAME DYNAMICS</b> . . . . .  | 77 |
| 5.1 Introduction, Notations and Preliminaries . . . . .  | 77 |
| 5.2 Initial Configuration . . . . .  | 80 |
| 5.3 Development of Temporal Model via Markov Chain . . . . .   | 81 |
| 5.4 Development of Spatial Dependence Model . . . . .  | 83 |

|   |   |     |
|---|---|-----|
| 5.5   | Existence of Joint Distribution of Spatial-temporal Model . . . . . | 84  |
| 5.6   | Future Work . . . . .   | 87  |
| <b>CHAPTER 6. OPTIMALITY AND STABILITY OF SYMMETRIC EVOLUTIONARY GAMES WITH APPLICATIONS IN GENETIC SELECTION . . . . .</b> |   | 88  |
| 6.1   | Introduction . . . . .  | 88  |
| 6.1.1   | Further Mathematical Background . . . . .                           | 90  |
| 6.1.2   | Further Biological Background . . . . .                             | 91  |
| 6.2   | GKproblems vs. SEgames . . . . .                                    | 92  |
| 6.3   | Second-order Optimality Conditions . . . . .                        | 98  |
| 6.4   | Evolutionarily Stable States . . . . .                              | 103 |
| 6.5   | Second-order Stability Conditions . . . . .                         | 104 |
| 6.6   | Computational Methods . . . . .                                     | 106 |
| 6.7   | Games for Genetic Selection . . . . .                               | 109 |
| 6.8   | Concluding Remarks . . . . .  | 114 |
| <b>CHAPTER 7. OPTIMAL STRATEGIES FOR SOCIAL NETWORK GAMES . . . . .</b>   |   | 117 |
| 7.1   | Introduction . . . . .  | 117 |
| 7.2   | Social Network Games . . . . .                                      | 118 |
| 7.3   | Social Network Games as Symmetric Evolutionary Games . . . . .      | 120 |
| 7.3.1   | Optimality Conditions . . . . .                                     | 121 |
| 7.3.2   | Stability Conditions . . . . .                                      | 124 |
| 7.4   | Cliques as Equilibrium States . . . . .                             | 125 |
| 7.4.1   | Is a Clique an Equilibrium State? . . . . .                         | 125 |
| 7.4.2   | Is a Maximal Clique an Equilibrium State? . . . . .                 | 127 |
| 7.4.3   | Is a Maximal Clique a Local Maximizer? . . . . .                    | 128 |
| 7.4.4   | Is a Maximal Clique a Strict Local Maximizer? . . . . .             | 133 |
| 7.4.5   | Is a Maximum Clique an Optimal and Stable Strategy? . . . . .       | 135 |
| 7.4.6   | Non-Clique Equilibrium States . . . . .                             | 138 |
| 7.5   | Conclusion . . . . .  | 140 |

|  |            |
|--|------------|
| <b>CHAPTER 8. SUMMARY AND CONCLUSION . . . . .</b>   | <b>146</b> |
| 8.1 2D Yeast Cooperation Game . . . . .  | 146        |
| 8.2 ESS in Nonlinear Game and the Asymptotic Behavior of Solution to Adapted Replicator Diffusion Equation . . . . . | 148        |
| 8.3 Parameter Estimation in Dynamical Metagenomic Model . . . . .  | 149        |
| 8.4 Markov Random Field Model for Analysis of Game Dynamics . . . . .  | 151        |
| 8.5 Optimality and Stability of Symmetric Evolutionary Games with Applications in Genetic Selection . . . . .        | 151        |
| 8.6 Optimal Strategies for Social Network Games . . . . .  | 153        |
| <b>APPENDIX A. PROOF OF THEOREMS . . . . .</b>   | <b>157</b> |
| <b>APPENDIX B. LOGIT TRANSFORMATION IN EM ALGORITHM . . . . .</b>  | <b>160</b> |
| <b>APPENDIX C. INNER-OUTER ITERATIVE PROCEDURE IN EM ALGORITHM . . . . .</b>   | <b>165</b> |
| <b>BIBLIOGRAPHY . . . . .</b>  | <b>167</b> |
| <b>PUBLICATION LIST . . . . .</b>  | <b>177</b> |

## LIST OF TABLES

## LIST OF FIGURES

|            |   |    |
|------------|---|----|
| Figure 2.1 | Design of the yeast experiment and sucrose metabolism in yeast. . . . .   | 7  |
| Figure 2.2 | Theoretical and experimental summary of cooperation in yeast system. (a) A model of cooperation with experimentally measured concave benefits. (b) Fraction of cooperators at equilibrium as a function of glucose and histidine concentrations. . . . .  | 8  |
| Figure 2.3 | The payoff functions with respect to cooperator frequencies and corresponding equilibrium for three types of matrix games. . . . .  | 11 |
| Figure 2.4 | The payoff functions with respect to cooperator frequencies and corresponding equilibrium for three types of nonlinear games. . . . .   | 14 |
| Figure 2.5 | The payoff functions with respect to cooperator frequencies and corresponding equilibria for the nonlinear games in yeast system. . . . .   | 14 |
| Figure 2.6 | Plots of average cooperator frequencies with respect to number of generations in two specific games with corresponding glucose and histidine parameters (1 corresponds to cooperator while 0 corresponds to defector). . . . .  | 16 |
| Figure 2.7 | Plots for log frequencies of cooperators using deterministic rules in the numerical experiments based on four types of initialized lattices (3 columns from left to right correspond to $N = 4, 8$ and $24$ ). (a) Randomized initial lattice. (b) Initial lattice with a square of cooperators in a population of defectors. (c) Initial lattice with a square of defectors in a population of cooperators. (d) Initial lattice with left side cooperators and right side defectors. . . . . | 18 |

|             |  |    |
|-------------|--|----|
| Figure 2.8  | Plots for log frequencies of cooperators using stochastic rules in the numerical experiments based on four types of initialized lattices (3 columns from left to right correspond to $N = 4, 8$ and $24$ ). (a) Randomized initial lattice. (b) Initial lattice with a square of cooperators in a population of defectors. (c) Initial lattice with a square of defectors in a population of cooperators. (d) Initial lattice with left side cooperators and right side defectors. . . . . | 20 |
| Figure 2.9  | Invasion conditions for defectors. . . . .   | 23 |
| Figure 2.10 | Invasion conditions for cooperators. . . . .   | 26 |
| Figure 3.1  | Two dimensional approximates for strategy $C$ frequency at different time points.<br>(a) $t = 0.2$ . (b) $t = 1.5$ . . . . .   | 49 |
| Figure 3.2  | Long time behavior of the average frequency of strategy $C$ in adapted replicator diffusion equation and comparison of it and replicator dynamics. (a) Replicator dynamics. (b) Adapted replicator diffusion equation. . . . .   | 49 |
| Figure 3.3  | Game theory models of cooperation in yeast system. (a) A model of cooperation with experimentally measured concave benefits yields a central region of parameter space that is a snowdrift game. (b) Fraction of cooperators at equilibrium as a function of the glucose and histidine concentrations. . . . .   | 51 |
| Figure 3.4  | Two dimensional approximates for the cooperator strain frequency using different parameter sets at different time points. (a) $A = 1, a = 1, g = 0.00, c = 0.0023, T = 5000, I = 20$ at $t = 0.2$ . (b) $A = 1, a = 1, g = 0.00, c = 0.0023, T = 5000, I = 20$ at $t = 1.5$ . (c) $A = 1, a = 1, g = 0.03, c = 0.0254, T = 5000, I = 20$ at $t = 0.2$ . (d) $A = 1, a = 1, g = 0.03, c = 0.0254, T = 5000, I = 20$ at $t = 1.5$ . . . . .  | 52 |

|            |  |    |
|------------|--|----|
| Figure 3.5 | Long time behavior of the average frequency of cooperators in adapted replicator diffusion equations and comparison of replicator dynamics and adapted replicator diffusion equations . (a)-(b) Snowdrift game where the parameter set is $A = 1, a = 1, g = 0.00, c = 0.0023, T = 5000, I = 20$ , and the experimental equilibrium is 0.6. (c)-(d) Prisoner's Dilemma where the parameter set is $A = 1, a = 1, g = 0.03, c = 0.0254, T = 5000, I = 20$ , and the experimental equilibrium is 0. . . . .  | 53 |
| Figure 4.1 | Observed temporal patterns for each OTU of 17 in each of 5 mice. . . . .   | 61 |
| Figure 4.2 | Frequencies of three groups solved by integrating the replicator equations at (a) the true parameters or (b) the parameters estimated during initialization. .   | 71 |
| Figure 4.3 | Simulated data for 17 OTUs of 5 mice at 21 time points. . . . .  | 71 |
| Figure 4.4 | Summary of estimated dispersion parameters. . . . .  | 72 |
| Figure 4.5 | Summary of estimates for parameters $s_{jm}$ and $d_{ik}$ . . . . .  | 73 |
| Figure 4.6 | Comparison of fitted exact solution with solution of differential equations. .   | 74 |
| Figure 4.7 | Standardized residuals for the fitting to the simulated data. . . . .  | 74 |
| Figure 5.1 | Illustration of set of inactive locations. . . . .   | 79 |
| Figure 6.1 | Two-Player Game: A two-player, two-strategy symmetric game is demonstrated. The strategies for player I are given in vector $x = (x_1, x_2)^T$ , and for player II in $y = (y_1, y_2)^T$ , $x, y \in S = \{x \in R^2 : \Sigma_i x_i = 1, x_i \geq 0, i = 1, 2\}$ . The fitness $A_{i,j}$ of strategy pair $(x_i, y_j)$ is given in the $(i, j)$ -entry of a $2 \times 2$ fitness matrix $A$ . A strategy pair $(x^*, y^*)$ is said to be optimal if $x^{*T} A y^* \geq x^T A y^*$ for all $x \in S$ and $y^{*T} A x^* \geq y^T A x^*$ for all $y \in S$ , when the game is said to reach the Nash equilibrium. . . . . | 90 |



- Figure 7.3 Social Cliques: In graph  $G$ , the nodes  $\{1, 2, 3\}$  form a clique. Let  $x^*$  be a strategy for choosing this clique,  $x_1^* = x_2^* = x_3^* = 1/3$  and  $x_i^* = 0$  for all  $i = 4, \dots, 8$ . Then,  $x^{*T}Ax^* = (Ax^*)_i = 2/3$  for all  $i$  such that  $x_i^* > 0$ , i.e.,  $i = 1, 2, 3$ . However,  $x^{*T}Ax^* \geq (Ax^*)_i$  does not hold for all  $i$  such that  $x_i^* = 0$ , i.e.,  $i = 4, \dots, 8$ , since  $x^{*T}Ax^* = 2/3 < (Ax^*)_4 = 1$ . By Theorem 7.3.1,  $x^*$  cannot be an optimal strategy for the SNGame on graph  $G$ . 126
- Figure 7.4 Maximal Cliques: The nodes  $\{1, 2, 3, 4\}$  in  $G$  form a maximal clique. Let  $x^*$  be a strategy on this clique,  $x_1^* = x_2^* = x_3^* = x_4^* = 1/4$  and  $x_i^* = 0$  for all  $i = 5, \dots, 8$ . Then,  $x^{*T}Ax^* = (Ax^*)_i = 3/4$  for all  $i$  such that  $x_i^* > 0$ , i.e.,  $i = 1, 2, 3, 4$ , and also,  $x^{*T}Ax^* \geq (Ax^*)_i$  for all  $i$  such that  $x_i^* = 0$ , i.e.,  $i = 5, \dots, 8$ . Then,  $x^*$  satisfies all the conditions in Theorem 7.3.1, and must be an optimal strategy for the SNGame on  $G$ . 127
- Figure 7.5 Non-Clique Strategies: The nodes  $\{1, 2, 3, 4, 5\}$  in  $G$  do not form a clique, but the strategy  $x^*$ ,  $x_1^* = x_3^* = x_4^* = 1/4$ ,  $x_2^* = x_5^* = 1/8$ , and  $x_i^* = 0$  for  $i = 6, 7, 8$ , is in fact optimal: It is easy to verify that  $x^{*T}Ax^* = (Ax^*)_i = 3/4$  for all  $i$  such that  $x_i^* > 0$ , i.e.,  $i = 1, 2, 3, 4, 5$ , and also,  $x^{*T}Ax^* \geq (Ax^*)_i$  for all  $i$  such that  $x_i^* = 0$ , i.e.,  $i = 6, 7, 8$ . Then,  $x^*$  satisfies all the conditions in Theorem 7.3.1, and must be an optimal strategy for the SNGame on  $G$ . 128
- Figure 7.6 Maximal Cliques vs. Local Maximizers: The nodes  $\{1, 2, 3, 4\}$  in  $G$  form a maximal clique. It is an equilibrium state for the SNGame on  $G$ , but not a local maximizer of the corresponding SPproblem: Let  $x^*$  be the equilibrium state on this clique,  $x_i^* = 1/4$  for all  $i = 1, \dots, 4$  and  $x_i^* = 0$  for all  $i = 5, \dots, 8$ . Construct a new state  $x = x^* + p$ , where  $p = (-2\epsilon, 0, 0, 0, \epsilon, \epsilon, 0, 0)^T$  for a small  $\epsilon > 0$ . Then, it is easy to verify that  $x_i \geq 0$  for all  $i = 1, \dots, 8$ ,  $\sum_i x_i = 1$ , and  $x^T Ax = x^{*T}Ax^* + 2\epsilon^2 > x^{*T}Ax^*$  for all small  $\epsilon > 0$ . As  $\epsilon$  goes to zero,  $x$  is arbitrarily close to  $x^*$ , yet  $x^T Ax > x^{*T}Ax^*$ . Therefore,  $x^*$  cannot be a local maximizer for the corresponding SPproblem. 129

Figure 7.7 Strict Local Maximizer: The nodes  $\{1, 2, 3, 4\}$  in  $G$  form a maximal clique.

The strategy  $x^*$  on this clique for the SNgame on  $G$ ,  $x_i^* = 1/4$  for all  $i = 1, \dots, 4$  and  $x_i^* = 0$  for all  $i = 5, \dots, 8$ , is a local maximizer of the corresponding SPproblem, and  $x^{*T}Ax^* \geq x^TAx$  for any  $x \in S$  in a small neighborhood  $U$  of  $x^*$ . However, if we choose  $x \neq x^*$  such that  $x_1 = x_3 = x_4 = 1/4$ ,  $x_2 = 1/4 - \epsilon$ ,  $x_5 = \epsilon$ , and  $x_i = 0$  for  $i = 6, 7, 8$ , we see for any  $U$  that  $x \in S \cap U$  for sufficiently small  $\epsilon > 0$ , and  $x^TAx = x^{*T}Ax^* = 3/4$ .

Therefore,  $x^*$  is not a strict local maximizer for the SPproblem. . . . . 133

Figure 7.8 Recovering Maximum Cliques: In the graph on top, the nodes  $\{1, 2, 3, 4, 5, 6\}$

form a subgraph  $H$ , and  $x^* \in S$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , is a global maximizer for the SPproblem. However,  $H$  is not a maximum clique. Since node 2 and 6 are not connected, we therefore add  $1/8$  to  $x_2$  but subtract it from  $x_6$ . We then obtain a reduced subgraph  $H$ , with  $V_H = \{1, 2, 3, 4, 5\}$ ,  $E_H = \{(i, j) \in E : i, j \in V_H\}$ , as shown in the second graph to the top. The solution  $x^* \in S$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , remains to be a global maximizer for the SPproblem. Next, since node 1 and 5 are still not connected, we then add  $1/8$  to  $x_1$  but subtract it from  $x_5$  to obtain a further reduced subgraph  $H = (V_H, E_H)$ ,  $V_H = \{1, 2, 3, 4\}$ ,  $E_H = \{(i, j) : i, j \in V_H\}$ , as shown in the graph in the bottom. The solution  $x^* \in S$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , again remains to be a global maximizer for the SPproblem, but this time,  $H$  is a maximum clique of  $G$ . . . . . 144

Figure 7.9 Non-Clique Equilibrium States: One of the equilibrium states of the SNgame

on graph  $G$  in the figure is when the population is distributed on the two disconnected cliques, with  $x_i^* = 1/8$  for all  $i = 1, \dots, 8$ . It is easy to verify that this strategy is not a local maximizer of the corresponding SPproblem and thereby, is not even weakly evolutionarily stable. . . . . 145

## ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude to those who helped me with various aspects of conducting research and the writing of this dissertation. First and foremost, Dr. Zhijun Wu and Dr. Karin Dorman, for their ceaseless motivation, encouragement, guidance and help during my Ph.D study and research. Without their support, I would not have been able to complete this work. With their motivation and encouragement, I started to explore the scientific curiosity deep in my mind and build mental toughness towards the difficulties. With their guidance and help, I started to learn the fun of research and spirit of collaboration.

I would like to thank Dr. Stephen Willson, Dr. Dan Nordman, Dr. Alexander Roitershtein and Dr. Wen Zhou, who have given me generous help and encouragement. Dr. Willson has given helpful suggestions and feedback on the research projects in some situations. Dr. Nordman has taught me the fun of probability and statistics in the theory course, and it has been an educational and interesting experience to collaborate with him. Dr. Roitershtein has provided nice recommendations and shown me how to maintain the attitude of working hard everyday. Dr. Zhou has collaborated with me on a few research projects, and has given me invaluable help and advices on research and life as an elder friend.

I am given the opportunity to pursue the Ph.D. degree by the Department of Mathematics and Department of Statistics, Iowa State University, to whom I owe my sincere gratitude. I would like to thank all my friends for supporting me and spending the six years with me. I also want to thank all the mentors who selflessly help me learn more efficiently and enrich my life along the way with optimism, confidence and hope.

## ABSTRACT

Modeling of evolution is becoming increasingly important in biological and social systems. The idea of evolution presents a reasonable and convincible perspective to the problem of long term competition between different species or populations in nature. Evolutionary game theory has seen tremendous achievements and it exhibits good performance in modeling competition of species or populations. However, there remain many challenges, and the underlying scientific mechanism of competition is full of uncertainties. Bringing tools of math-biology and statistics to modeling the competitive phenomena provides a way of throwing off the veil of nature to reveal itself to the world.

In this thesis, we study the spatial effects on evolution in game models numerically and theoretically, optimality and stability of symmetric evolutionary game and its applications to genetic selection and social network, and parameter estimation in game theoretic modeling of the biological and social systems. More precisely, we categorize the types of nonlinear games and investigate the simulation study of spatial effects on the evolution of cooperation in a nonlinear yeast game. We also incorporate the spatial diffusion effects in the replicator dynamics of nonlinear game models, and prove the asymptotic behavior of the solution to the corresponding adapted replicator diffusion equation. Furthermore, we apply the statistical techniques and methodologies to solve the inverse problem in evolutionary game dynamics, that is, the NB mixture model and Markov random field model incorporating replicator equations are built where penalized approximated negative log-likelihood method with generalized smoothing approach and Besag pseudo-likelihood method are implemented to facilitate the estimation and inference of model parameters. Finally, the theory for obtaining optimal and stable strategies for symmetric evolutionary games is explored, and new proofs and computational methods are provided. And the symmetric evolutionary game is applied to model the evolution of a population over a social network, then several different types of equilibrium states corresponding to social cliques are analyzed and a set of conditions for their stabilities are proved.

## CHAPTER 1. INTRODUCTION

Evolutionary game theory is an important and comprehensive way to look at the world. It arises whenever the fitness of an individual is not constant but depends on the relative abundance of others in the population. And in evolutionary game theory, the replicator equation is an essential differential equation that describes frequency-dependent selection among a fixed number of species or strategies. Evolutionary game theory and ecology are linked via the equivalence between the replicator equation and Lotka-Volterra equation of ecological systems. The main ingredients of evolutionary game theory are reproduction, selection and spatial movement. And population is the fundamental basis of every evolution. Individuals, genes, or ideas change over time, and population evolve.

### 1.1 Evolutionary Game Theory

#### 1.1.1 Evolutionary Theory

The original formulation of evolutionary theory dealt with the genetic evolution of the origin and adaptation of the species in nature. But more recently the idea of evolution has expanded to all fields including biology and related disciplines of the life science. Wherever information reproduces, there is evolution. Mutations are caused by errors in information transfer, resulting in different types of messages. Selection among types emerges when some messages reproduce faster than others. Mutation and selection make evolution. Mutation and selection can be described by a lot of mathematical equations and statistical models. Therefore evolution has become a theory based on mathematics and statistics.

Evolutionary theory presents the principles according to which life has evolved and continues to evolve. The study of evolution has grown enormously since the 1950s. It is driven by the quest to understand the world we live in and the stuff we are made of. All observation of a living system could ultimately be interpreted in the context of its evolution. Darwin's evolutionary theory has provided

a general guideline or principle for how life is developed and evolved, but the question now is that can we add bricks to build a solid foundation for it, in other words, can we make our contribution to the sophisticated system of mathematical equations and statistical models to support, explain, or even extend the evolutionary theory.

### 1.1.2 Game Theory

Game theory was established by John von Neuman and Oskar Morgenstern. They wished to build a mathematical theory, for example, using quantitative models of competition and cooperation between the decision makers, to study human behavior in strategic and economic decisions. Later John Nash participated and created a simple but important concept in game theory, which is now called Nash equilibrium. Another type of equilibrium named evolutionarily stable strategy (ESS) is very similar to Nash equilibrium. Both concepts are important for game theory. William Hamilton and Robert Trivers were among the first to use game theoretical ideas in biology. John Maynard Smith and George Price introduced game theory to evolutionary biology and population thinking to game theory. The field of evolutionary game theory was also founded by the work of Peter Taylor, Josef Hofbauer, Karl Sigmund and other scientists.

Traditional game theory typically analyzes an interaction between two players and deals with the problem, for example, how each player can maximize her or his payoff in a game given that each player do not know what the other player do. There are four basic elements in a game: players, information, strategies and payoffs. And the concept of rationality are usually assumed in a game for the individuals come to play. However, evolutionary game theory does not necessarily reply on rationality. It considers a population of players interacting in a game. Individuals have fixed strategies and interact randomly with each other. Payoff is then interpreted as fitness, and success in the game is translated into reproductive success. Strategies that do well reproduce faster. This is straightforward natural selection.

## 1.2 Outline of the Thesis

In this thesis, we bridge the areas of evolutionary theory, game theory with population dynamics, and mathematical and statistical modeling in biological and life science systems. The overarching

goal of the research is to study the spatial effects on evolution of models, optimality and stability of symmetric game and its applications, and parameter estimation in game theoretic modeling of the biological and social systems. We categorize the types of nonlinear games and investigate the effect of spatial structure on the evolution of cooperation in a nonlinear yeast game via simulation studies. We also incorporate the spatial diffusion term in the replicator dynamics of nonlinear game to obtain the adapted replicator diffusion equation, and prove the asymptotic behavior of the solution to this adapted parabolic equation. Furthermore, we apply the statistical techniques and methodologies to solve the inverse problem in evolutionary game dynamics, that is, the NB mixture model and Markov random field model incorporating replicator dynamics are built where penalized approximated negative log-likelihood method with generalized smoothing approach and Besag pseudo-likelihood method are implemented to facilitate the estimation and inference of model parameters. Finally, the theory for obtaining optimal and stable strategies for symmetric evolutionary games is explored, and new proofs and computational methods are provided. Also the symmetric evolutionary game is applied to model the evolution of a population over a social network, and several different types of equilibrium states corresponding to social cliques are analyzed with a set of conditions for their stabilities proved.

The outline of the thesis is the following. First of all, in Chapter 2, we derive the conditions to distinguish several types of nonlinear games, and investigate the effect of spatial structure on the evolution of cooperation in the nonlinear yeast game computationally. In Chapter 3, we study the equivalent conditions to evolutionarily stable strategies in general nonlinear games, and prove the long term behavior of the solution to the replicator diffusion equation adapted from the replicator dynamics in nonlinear games. We then investigate the famous inverse problem by using statistical models and techniques in Chapter 4 and 5. In particular, in Chapter 4 we establish a NB mixture model with generalized smoothing approach to study the interaction among various OTUs in microbial communities, which is expected to be useful for a variety of evolutionary game studies. In Chapter 6, we review the theory for obtaining optimal and stable strategies for symmetric evolutionary games, and provide new proofs and computational methods, which is applied to justifying the strategies and their stabilities for a special class of genetic selection games. In Chapter 7, the symmetric evolutionary game is defined over a social network, and the correspondence between equilibrium states and social cliques are proved. Finally we sum up and conclude in Chapter 8.

## CHAPTER 2. 2D YEAST COOPERATION GAME

A paper to be submitted

Min Wang, Yuanyuan Huang, Zhijun Wu

### Abstract

Evolutionary game theory plays an important role in cooperation problem of evolutionary biology, and there are mainly two games that scientists are enamored by in recent decades: the Prisoner’s Dilemma and the Snowdrift Game. Numerous experimental studies have demonstrated the distinct evolutionary dynamics of cooperative behavior in these two linear games from intriguing explorations. We generalize the study on these two simple linear games to the nonlinear games, and investigate the effect of spatial structure on evolution of cooperation computationally. We apply the numerical procedure to a specific nonlinear game raised by Gore et al. who studied the coexistence of wildtype cooperator strain and mutant cheater strain in a yeast system. Along with the simulation procedure, we establish the dynamic difference equations for the simulation study which boil down to replicator equations under certain conditions. Finally we analyze the conditions for two circumstances of invasion which are consistent with the simulation studies. We expect that the proposed analysis and conclusions will be widely applied to a variety of nonlinear games.

### 2.1 Introduction

The phenomenon of cooperative interactions among microbial species has attracted scientists intuitively over several decades. A lot of experimental results have proved the persistence of cooperation and coexistence which is closely related with spatial scales in the competition environment. Kerr, Riley, Feldman and Bohannan (2002) experimentally tested the biodiversity of colicin-producing strain (C),

sensitive strain (S) and resistant strain (R) in a C-S-R model, and proposed that these three strains with non-hierarchical relationship would coexist when dispersal and interaction were local but the diversity was rapidly lost when dispersal and interaction occurred over large spatial scales in (10). Griffin, West and Buckling (2004) explored the interaction between relatedness and the spatial scale of competition and found that the scale competition was of particular importance to the evolution of cooperation in microorganisms (6).

The exact relationship between the coexistence and the spatial scale has also been numerically verified by immense simulations. Nowak (1992) mentioned that the generated spatial structure within populations may often be crucial for the evolution of cooperation in (12). The evolvement of mutualism that required spatial structure in competitive interactions was established by Doebeli and Knowlton (3). Durrett (1997) analyzed the stability of the spatially structured population given set of parameters and concluded that coexistence was easily achievable in a spatially explicit system (4). However Hauert and Doebeli (2004) showed the property that spatial structure promoted the evolution of cooperation in Prisoner's Dilemma did not work for Snowdrift Game anymore and instead the spatial structure reduced the proportion of cooperators in (8).

We introduce a yeast game Gore et al. proposed at 2009. They manipulated the microbial interactions in models of cooperation and competition between two yeast strains. One type of yeast was the wild-type cooperator who produced invertase, hydrolyzed sucrose and released the glucose and fructose, of which vast majority diffused away and the rest small proportion was imported into the cell. The other type was the mutant cheater that did not produce invertase but instead took advantage of and invaded a population of wild-type cooperator cells. Gore et al. fitted the experiment results with the nonlinear defector's payoff  $P_D$  and cooperator's payoff  $P_C$ . And both payoff functions were defined with respect to the frequencies of cooperators, and changed over a set of parameters including the glucose and histidine concentrations. Also they found that the competition was Prisoner's Dilemma when the concentration level of the histidine was low and that of the glucose was high, and the game switched to Snowdrift Game if the concentration of histidine was high and that of glucose was low. Based on these, we examined the mathematical properties of this model with varying experimental parameters in (9). We also simulated the growths of the yeast trains and compared them with the experimental results. Finally the equilibrium states of the system were computed and analyzed.

In all of the two-strategy games with possible numerical simulations, scientists have considered the two most famous ones: Prisoner's Dilemma and Snowdrift Game, and these two games are usually formulated in linear or matrix forms since the payoff functions of two strategies are assumed to be linear functions of the frequencies or equivalently functions of matrix-multiplied-vector forms. Here the generalization of these two games are introduced and defined in the sense that the payoff functions can be explicitly expressed by nonlinear functions of frequencies and the relationship between the two strategies could be quantified. From the nonlinear payoff functions, we analyze the mechanism of cooperation on a popular two dimensional lattice and experimentally simulate the interactions among individuals on the lattice. In the simulations, we gradually increase the neighborhood size of interaction which is a control parameter and appropriately describes the spatial effects. Then from the simulated investigations with corresponding neighborhood size, we sum up the conclusion that there is a strong relationship between the strength of local cooperation and spatial parameter, and as the neighborhood size increase, the frequency of cooperation is decreasing and eventually converges to the equilibrium frequency that is observed in the experiment.

From the simulation study that evolution of cooperation is closely related with the spatial structure and the convergence of the cooperation frequency to experimental equilibrium as interaction region extends to the overall population, we are able to build corresponding theoretical equation to interpret the underlying mechanism. It's known that the general continuous form of the differential equation associated with evolutionary game theory is the replicator dynamics equation. However the continuous time and space involved in the derivative is a subtle issue when we formulate the simulation of discrete process which is implemented from the idea of replicator equation. That is, all updates in the simulation are applied at both sequential time points and randomized grids which characterize the temporal and spatial discreteness. When we consider the changing interaction neighborhood, in other words, tuning the spatial effect, it is equivalent to incorporate and change the spatial parameter in the simulated process. So in order to set up the mathematical theory, it's natural to modify the common replicator equation which yields to a simulation-based replicator equation with the spatial parameter included. Finally from the modified replicator equation, we analyze the mechanism of local cooperation and the invasion conditions for single cooperator and single defector, which explain the numerical simulations well.

The rest of the paper is organized as follows. Section 2.2 introduces the yeast system and corresponding game dynamics. In section 2.3, we describe the traditional linear games and particular simulation explorations. Sections 2.4 generates the defined concept of traditional games to nonlinear games and presents the interesting simulation study of 2D yeast game. In section 2.5 we propose the difference equation to interpret the simulation and analyze the invasion conditions for two circumstances. Section 2.6 concludes.

## 2.2 Yeast System and the Nonlinear Payoffs

The interesting game between two strains—the wildtype cooperator strain (C) and the mutant cheater or defector strain (D) in a yeast system is undertaken by Gore et al., and the strains are manipulated as a microbial interaction in ways which have leaded to microbial models of cooperation and competition. As described in Figure 2.1, sucrose is often used as a carbon sources for yeast and it is hydrolyzed by the enzyme invertase which creates glucose and fructose. The majority of these monosaccharaides lose away before they are imported in to the cell which leads to cooperative sucrose metabolism that is exploited by the mutant cheater strain. The main difference between the wildtype cooperator strain and the mutant cheater strain is that the cooperator has an intact SUC2 gene which encodes invertase used in the hydrolysis while the cheater lacks the SUC2 gene. Also cooperator is histidine auxotroph since it has a defective HIS3 gene but the cheater possesses an intact HIS3 gene, and this difference imposes an extra cost on the cooperator.

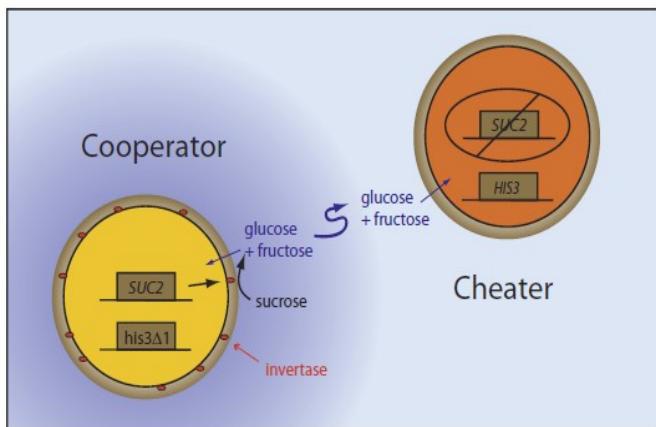


Figure 2.1 Design of the yeast experiment and sucrose metabolism in yeast.

Gore et al. experimentally investigated the coexistence between the cooperator and cheater strains and they found that the nature of interaction between these two strains or two strategies is the so-called Snowdrift Game or Hawk-dove game. The Snowdrift Game derives its game from the potentially cooperative interaction present when two drivers are trapped behind a pile of snow, and each driver must decide whether to clear a path. In this game, the optimal strategy is the opposite of the opponents, that is, cooperate when your opponent defects and defect when your opponent cooperates. The Snowdrift Game is qualitatively different from the Prisoner's Dilemma, in which all players have the incentive to cheat regardless of the strategies being followed by the others. In the experiment, the growth rate is measured as a function of glucose concentration and this function is actually concave and nonlinear. Furthermore the capture efficiency  $\epsilon$ , cost  $c$  and the power of  $\alpha = 0.15 \pm 0.01$  are imposed onto the benefit terms from the concavity of the observed benefits. Then the payoff functions of cheater or defector and cooperator are approximately the following:

$$P_D = [f(1 - \epsilon)]^\alpha$$

$$P_C = [\epsilon + f(1 - \epsilon)]^\alpha - c,$$

where  $f$  is the fixation or frequency of cooperators.

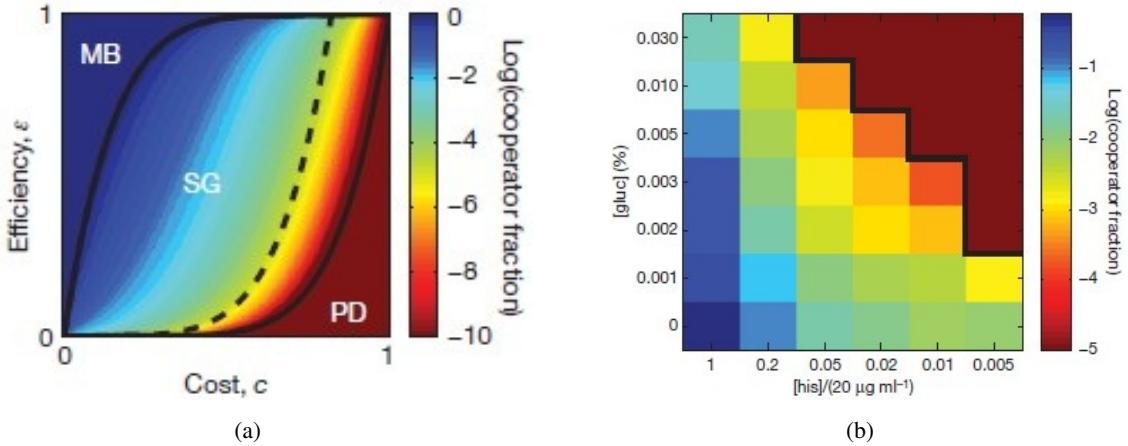


Figure 2.2 Theoretical and experimental summary of cooperation in yeast system. (a) A model of cooperation with experimentally measured concave benefits. (b) Fraction of cooperators at equilibrium as a function of glucose and histidine concentrations.

In Figure 2.2, we notice that varying the glucose concentration and histidine transforms the outcome of competition, that is, the type of the yeast game changes with respect to different parameters. More

specifically, the interaction between the cooperator and defector is a Mutual Beneficial game when the histidine level is very high and the glucose concentration is almost zero, a Prisoner's Dilemma game when the histidine level is low and the glucose concentration is large and a Snowdrift Game otherwise. There are different equilibrium states corresponding to different games. Furthermore, the equilibrium states are also different even for the same type of game but different parameters. Therefore the parameters are key elements in the nonlinear games and different subsets of full parameter space may change both the type and equilibrium state of the games.

### 2.3 Linear Games and Simulations

We consider a two-strategy evolutionary game defined by the matrix  $A$  which is called the fitness or payoff matrix in Table 2.1. Strategy C is also denoted as strategy 1 and strategy D corresponds to strategy 2. Let  $S = \{x \in R^2 : x \geq 0, \sum_{i=1}^2 x_i = 1\}$  be the set of all mixed strategies. Denote the strategy played by player 1 be  $x \in S$  and strategy played by player 2 be  $y \in S$ . The strategy for each player could be a pure strategy type, i.e.,  $e_i$  for  $i \in \{1, 2\}$  where  $e_i$  is the  $i$ th unit vector, or a mixed one with no element equal 1.

Table 2.1 Payoff or fitness matrix for linear game with two strategies  $C$  and  $D$

|          |            | Player 2   |            |
|----------|------------|------------|------------|
|          |            | Strategy C | Strategy D |
| Player 1 | Strategy C | a          | b          |
|          | Strategy D | c          | d          |

For the pure strategy type games, there are four possible paired strategy choices for the players, that is,  $(C, C)$ ,  $(C, D)$ ,  $(D, C)$  and  $(D, D)$ , where the first coordinate is the strategy for player 1 and the second for player 2. And the elements  $a, b, c$  and  $d$  in matrix  $A$  characterize the fitness of player 1 playing with player 2 for those choices. That is, player 1 with strategy  $C$  gets payoff  $a$  when playing against play 2 with strategy  $C$ ; player 1 with strategy  $C$  gets payoff  $b$  when playing against play 2 with strategy  $D$ ; player 1 with strategy  $D$  gets payoff  $c$  when playing against play 2 with strategy  $C$ ; player 1 with strategy  $D$  gets payoff  $d$  when playing against play 2 with strategy  $D$ . On the other hand, in the mixed strategy type games, a mixed strategy  $s \in S$  can be interpreted in the sense that the player choose pure strategy  $i$  with probability  $s_i$  for  $i \in \{1, 2\}$ . More specifically, if the pure strategy type game is repeated

over large number of times, then the frequency for a typical player to play strategy  $i$  is  $s_i$ . Denote the strategy profile for player 1 to be  $x$  and for player 2 be  $y$ , then the fitness for player 1 can be defined by  $\pi(x, y) = x^T A y$  and for player 2 by  $\pi(y, x) = y^T A x$  where  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ . If  $c > a > d > b$ , then this type of game is Prisoner's Dilemma game. If  $c > a > b > d$ , then it becomes the Snowdrift Game.

It is assumed that all players tend to maximize their fitness in competitions of games. If the game is Prisoner's Dilemma, each player will choose pure strategy D over all the possible strategies to obtain the optimal fitness. If the nature of the game is Snowdrift Game, both players will choose optimal strategy  $x^* = y^*$  where  $x_1^* = y_1^* = (b - d)/(b + c - a - d)$ ,  $x_2^* = y_2^* = (c - a)/(b + c - a - d)$ . In the optimization process, there is a phenomenon which actually reveals the cooperation and competition of two strategies. That is, strategy D always dominates strategy C in Prisoner's Dilemma. However, in Snowdrift Game, strategy C and strategy D are evenly matched and each takes over certain proportion of the population.

We extend the case of two players to a population of competing individuals with two strategies, and model the evolution of cooperation in the population. Denote the individuals choosing strategy C to be C players and the ones choosing strategy D be D players. Let  $f_C$  be the proportion of C players and  $f_D$  be that of D players, then  $(f_C, f_D)$  is the population profile. Let the expected payoff of typical C and D player be  $P_C$  and  $P_D$  respectively, then the average increase in fitness of C players is  $P_C = af_C + bf_D$  and that of D players is  $P_D = cf_C + df_D$ , which are called the payoff functions of C players and D players respectively and are both linear functions of the frequencies, from which the linear game derive its name. Note that a system of differential equations called replicator dynamics describe the evolution of the frequencies of these two strategies in a population and the underlying mechanisms can be modeled by the following 2-strategy differential equation on  $S$ :

$$\begin{aligned} \dot{f}_C &= f_C(P_C - \phi) = f_C(af_C + bf_D - \phi) \\ \dot{f}_D &= f_D(P_D - \phi) = f_D(cf_C + df_D - \phi), \end{aligned} \tag{2.1}$$

where the average payoff  $\phi$  is  $f_C P_C + f_D P_D = f_C(af_C + bf_D) + f_D(cf_C + df_D)$ . Since  $f_D = 1 - f_C$ , then system (2.1) is equivalent to

$$\dot{f}_C = f_C(1 - f_C)(P_C - P_D) = f_C(1 - f_C)[(a - b - c + d)f_C + b - d]. \tag{2.2}$$

The fixed points of equation (2.2) are  $f_C = 0, f_C = 1$ , and all the values of  $f_C \in (0, 1)$  satisfying  $(a - b - c + d)f_C + b - d = 0$ . If the game is Prisoner's Dilemma, then  $(a - b - c + d)f_C + b - d < 0$  for  $f_C \in (0, 1)$  which means  $\dot{f}_C < 0$  for  $f_C \in (0, 1)$ , and this implies  $f_C$  will decrease and finally reaches the equilibrium  $f_C = 0$ . If the game is Snowdrift Game, then there exists  $f_C^* = (b - d)/(b + c - a - d)$  such that  $(a - b - c + d)f_C + b - d = 0$ . Then  $\dot{f}_C > 0$  for  $f_C \in (0, f_C^*)$  and  $\dot{f}_C < 0$  for  $f_C \in (f_C^*, 1)$ , which indicates  $f_C$  increases on  $(0, f_C^*)$  and decreases on  $(f_C^*, 1)$ . So there is an interior equilibrium  $f_C^*$  for the Snowdrift Game. This scenario is called coexistence between C players and D players. If the game has elements satisfying  $a > c$  and  $b < d$ , then  $(a - b - c + d)f_C + b - d < 0$  for  $f_C \in (0, (d - b)/(b + c - a - d))$ , i.e.,  $\dot{f}_C < 0$  for  $f_C \in (0, (d - b)/(b + c - a - d))$ , and  $(a - b - c + d)f_C + b - d > 0$  for  $f_C \in ((d - b)/(b + c - a - d), 1)$ , i.e.,  $\dot{f}_C > 0$  for  $f_C \in ((d - b)/(b + c - a - d), 1)$ . Thus there are two Nash equilibria  $f_C = 0$  and  $f_C = 1$ . This is called bistability between strategies  $C$  and  $D$ . These three situations are illustrated in Figure 2.3.

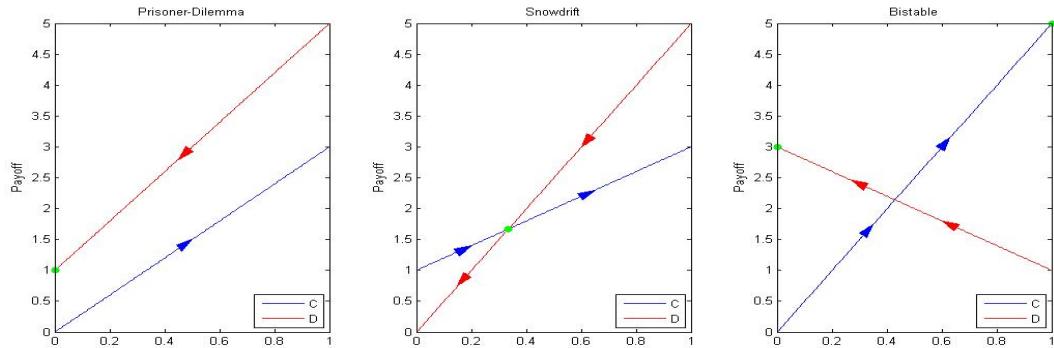


Figure 2.3 The payoff functions with respect to cooperator frequencies and corresponding equilibrium for three types of matrix games.

The three most common possibilities for the models of two strategies linear games are Prisoner's Dilemma, Snowdrift Game and Bistable-Equilibrium Game. Scientists prefer simulation studies on the first two types of game models. The mechanism underlying the modeling equation (2.2), which suggests the change in time dependent frequency of C players is proportional to the difference between the payoff of C players and that of D players, is of great importance to the simulation. And the differential equation models the frequency dependent selection in the deterministic limit of an infinitely large, well-mixed population, where stochasticity and spatial effects are ignored.

However, when the comparison between these two strategies is restricted to a local structure called neighborhood or interaction region where spatial effects are incorporated in simulations, it will correspondingly affect the center or focal individual whose strategy follows an update rule which is related with the replicator dynamics. These two-strategy game simulations are often established on a square lattice where each grid is occupied by an individual who plays as either a cooperator or a defector. Lattices are initialized randomly with certain proportions of the two strategies. The general assumption is that at each generation the current individuals at chosen update grids interact with their neighbors including themselves which yields payoff for these individuals respectively. And two typical neighborhoods are von-Neumann neighborhood which includes four nearest neighbors and Moore neighborhood with eight nearest neighbors. Similarly, for any of these two types of neighborhoods being used, each grid in the neighborhood region of updating center individual has a payoff, and once the payoffs of all the individuals in the neighborhood region are collected, the individual or strategy with the highest payoff invade the central player at next time point with a certain probability which is a function of the payoffs. Each time the update rule is closely related with the update probability. If the probability is always one, then the update is often called deterministic update; if the probability is less than or equal to one, then the update is called stochastic update. There are also two choices for the region of update grids, that is, two types of timekeeping updates—synchronized and asynchronous ones. Synchronized update correspond to populations with nonoverlapping generations and all individuals interact with all their neighbors and update their strategy in a synchronized fashion. Asynchronous or random update is that individuals are randomly selected and then interact with their neighbors, which yield the update of their strategy. Note that stochasticity of the simulation is introduced by both settings of asynchronous and stochastic updates.

Through the comparison of linear game simulations with results from mean field systems, some literature are able to pinpoint the effects solely arsing from the rigid spatial structure of the population (4; 7; 8). In particular, they have shown that for the Prisoner’s Dilemma, space can outweigh the inherent advantage of defectors over cooperators—at least for a small parameter range. For the Snowdrift game, a counter-intuitive result is obtained and justified several times that spatial structures tend to result in more frequent escalations in intra-species competitions.

## 2.4 Nonlinear Games and 2D Simulations

We generalize the selection dynamics as the nonlinear payoff functions for the two strategies  $C$  and  $D$  in population games are considered. That is, the replicator equations are

$$\begin{aligned}\dot{f}_C &= f_C(P_C - \phi) \\ \dot{f}_D &= f_D(P_D - \phi),\end{aligned}\tag{2.3}$$

where the average payoff  $\phi$  is  $f_C P_C + f_D P_D$  and at least one of  $P_C$  and  $P_D$  are nonlinear functions of the frequency  $f_C$  or  $f_D$ . Since  $f_D = 1 - f_C$ , then system (2.3) is equivalent to

$$\dot{f}_C = f_C(1 - f_C)(P_C - P_D).\tag{2.4}$$

Similar to the three types of games in linear games, we extend their definitions to nonlinear games with nonlinearity for payoff functions in terms of strategy frequencies. And the analytical equilibria for three type of nonlinear games are described by the plots in Figure 2.4. The comparative relationship between  $P_C$  and  $P_D$  in nonlinear games is almost the same as that in linear games showed in Figure 2.4 except the skewness and convex or concave behavior in Snowdrift Game. From Figure 2.5, we can see that strategy  $D$  dominates strategy  $C$  in Prisoner's Dilemma where dominate means  $P_D > P_C$  for any  $f_C \in [0, 1]$ . For Snowdrift Game,  $P_C$  and  $P_D$  intersect at a point with  $f_C = f^*$  and strategy  $C$  dominates strategy  $D$  for  $f_C \in [0, f^*)$  while strategy  $D$  dominates strategy  $C$  for  $f_C \in (f^*, 1]$  and this indicates  $f_C = f^*$  is an interior Nash equilibrium. However the third plot suggests a completely different situation from the Snowdrift Game. There is also an intersection point  $\hat{f}$  between  $P_C$  and  $P_D$ . However,  $P_C > P_D$  for  $f_C \in (\hat{f}, 1]$ , and  $P_D > P_C$  for  $f_C \in [0, \hat{f})$ , which indicates there are two isolated Nash equilibria  $f_C = 0$  and  $f_D = 1$ . So the nonlinear system illustrated by this plot is bistable for both strategies  $C$  and  $D$ .

For the nonlinear payoff functions in yeast system that Gore et al. have proposed in their paper, the more specific forms are:

$$\begin{aligned}P_D &= [g + (1 - g)f(1 - \epsilon)]^\alpha \\ P_C &= [g + (1 - g)(\epsilon + f(1 - \epsilon))]^\alpha - c,\end{aligned}$$

where  $f = f_C$  is the frequency of cooperators and  $g, c$  are parameters correspond to glucose concentrations and cost. By investigating the pattern and property of the nonlinear payoff functions for two

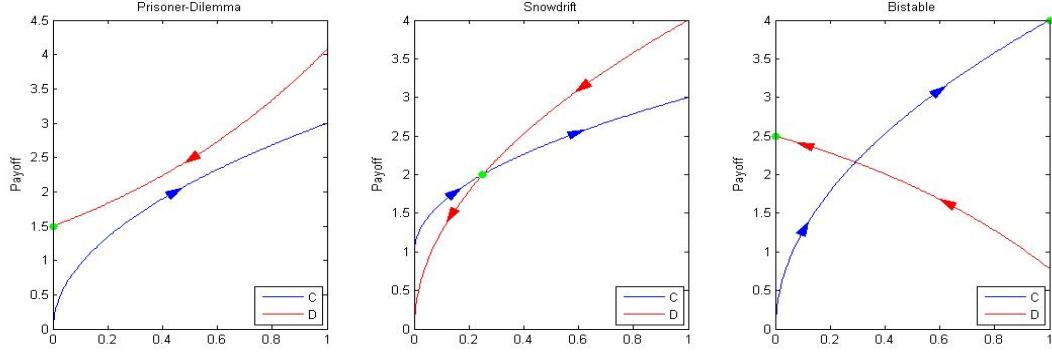


Figure 2.4 The payoff functions with respect to cooperator frequencies and corresponding equilibrium for three types of nonlinear games.

strategies C and D, we observe that there are essentially two types of nonlinear games—Prisoner's Dilemma and Snowdrift Game for different combinations of parameters  $g$  and  $c$  in this system, and two typical examples of these are shown in Figure 2.5.

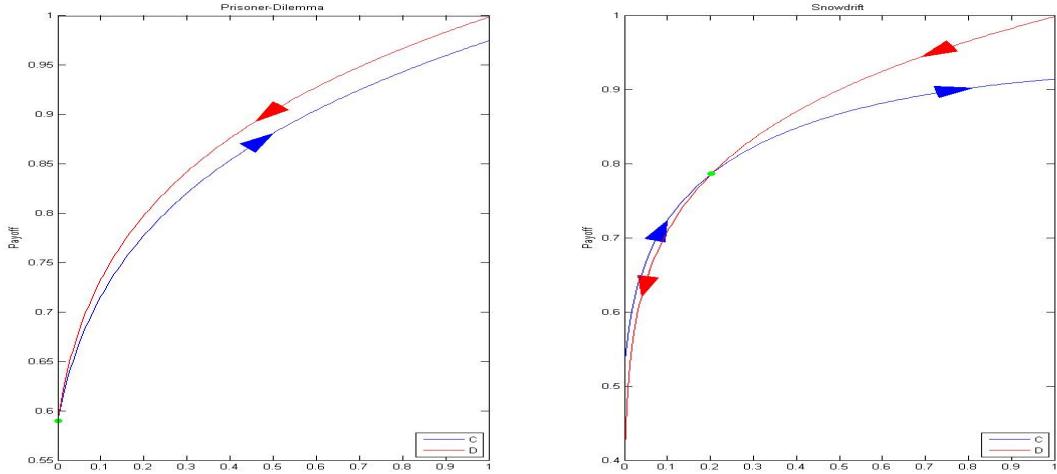


Figure 2.5 The payoff functions with respect to cooperator frequencies and corresponding equilibria for the nonlinear games in yeast system.

By intuitively obtaining the idea from the selection dynamics or replicator equation  $\dot{f}_C = f_C(1 - f_C)(P_C - P_D)$ , we use the nonlinear payoff functions in yeast game to implement the spatial simulation with different neighborhood sizes  $N$  which will be described.

The procedure of 2D simulation is as followed. First of all, the spatial structured population is set up, and individuals are confined to sites on a regular  $100 \times 100$  lattice with periodic boundary conditions which means the edges are wrapped around to generate a toroidal universe such that all of them interact with their neighbors. The neighborhood or interaction region is tuned with parameters  $N = 4, 8$  and  $24$ . Then the square lattice is initiated, and each individual is randomly assigned with either strategy C or D, or specifically assigned with some particular patterns of the two strategies. Note here we only consider pure strategy for each individual. Then each time we randomly pick the interaction region which consists of one grid, and compute the nonlinear payoff of each neighbor which depends on the frequency of C in the neighborhood of each neighbor and finally yields to the averaged payoff for C and D. Afterwards we update the center individuals based on well-defined update rules. There are empirically two rules to update the center. One is the deterministic rule and the other is the stochastic rule. The deterministic rule is the revised best-take-over rule according to which the strategy with the highest average payoff will take over the center individual with probability 1 while the generic best-take-over rule is that the individual with the highest payoff in a given neighborhood replacing the center. The stochastic rule is that the strategy with the highest average payoff will replace the center individuals with certain probability which depends on the difference of the average payoff difference of two strategies. One type of probability we use is the form of expected mean in logistic distribution with one parameter determining the scale. Finally the results of the equilibrium levels are determined by evolving the lattice over  $M$  ( $M = 1,000,000$ ) generations and then averaging the cooperator strategies over another 1,000 generations. They will show that the frequency of cooperators converges to a stable state as the running time goes to infinity.

Table 2.2 Frequencies of cooperator at equilibrium for various glucose and histidine

| Glucose<br>Histidine \ | 0      | 0.0010 | 0.0020 | 0.0030 | 0.0050 | 0.0100 | 0.0300 |
|------------------------|--------|--------|--------|--------|--------|--------|--------|
| 0.0050                 | 0.0082 | 0.0012 | 0      | 0      | 0      | 0      | 0      |
| 0.0100                 | 0.0071 | 0.0046 | 0.0006 | 0.0001 | 0      | 0      | 0      |
| 0.0200                 | 0.0122 | 0.0051 | 0.0010 | 0.0007 | 0.0002 | 0      | 0      |
| 0.0500                 | 0.0208 | 0.0112 | 0.0032 | 0.0017 | 0.0010 | 0.0004 | 0      |
| 0.2000                 | 0.1097 | 0.0643 | 0.0201 | 0.0116 | 0.0060 | 0.0038 | 0.0015 |
| 1.0000                 | 0.6000 | 0.2950 | 0.2025 | 0.1825 | 0.1000 | 0.0400 | 0.0250 |

The data Gore et al. collected in their experiment is displayed in Table 2.2. The table shows us that there are 42 observed equilibria associated with 42 combinations of glucose and histidine level which correspond to 42 choices of parameter settings including the cost  $c$ . We implement the numerical simulation with neighborhood sizes  $N = 4, 8, 24$  and different initialized matrices for the lattice. In the process of simulation, there are two types of update rules—deterministic and stochastic rules used. Deterministic rule means each individual will certainly choose the strategy that posses higher average payoff in its neighborhood region at each generation while stochastic rules means the strategy with higher average payoff will take the focal point with the probability which depends on the difference of average strategy payoffs. The computational results based on them are generally different with more details as followed.

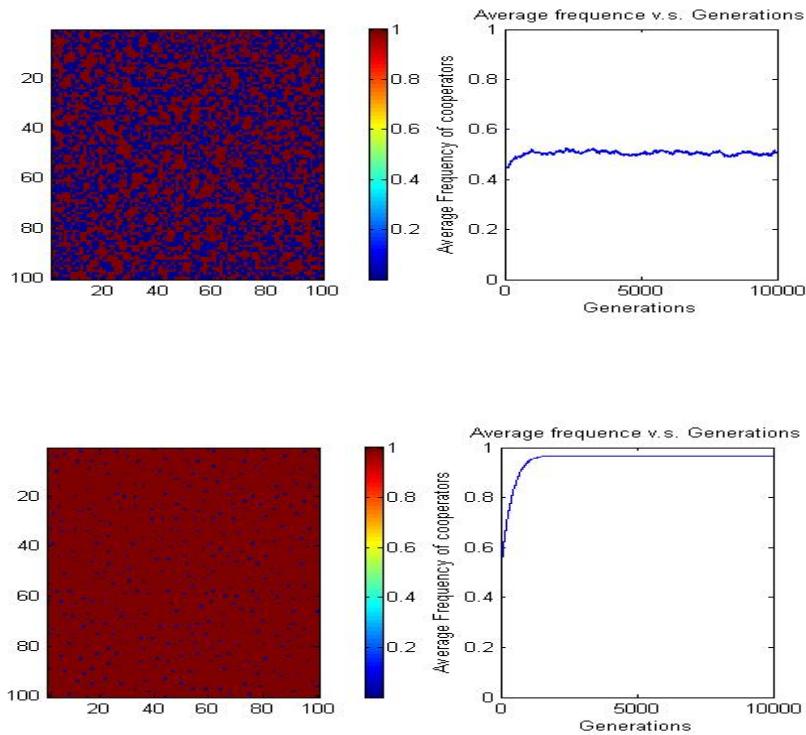


Figure 2.6 Plots of average cooperator frequencies with respect to number of generations in two specific games with corresponding glucose and histidine parameters (1 corresponds to cooperator while 0 corresponds to defector).

### 2.4.1 Deterministic Rule

The numerical experimental plots in Figure 2.6 show us that the average frequencies of cooperators over finite large  $M$  generations stay almost the same after certain number of generations. These curves of them also suggest that the average frequencies of cooperators for different games are all getting closer to the "approximate" true equilibria or more precisely some equilibria with small enough oscillations as the number of generations increases. The actual pattern of cooperators and defectors is changing all the time, and the frequency of cooperators, in a sufficiently large array, is almost constant, which is called a "dynamic equilibrium". But this approximation is not static since there are still minor wiggles around the "approximate" true equilibria. They suggest that the average frequency of cooperators stays at the small range close to some constant after  $M$  time steps of generation. So we collect the average frequencies of cooperators over certain generations after  $M$  generations, and the results for all cases are displayed in Figure 2.7.

For different lattices that are chosen specifically, the simulations are applied in three different neighborhood sizes respectively, and the frequencies of cooperation after  $M$  iterations are provided. Figure 2.7 suggests that for different initialized lattices, the final dynamic equilibrium of the cooperator frequencies follows almost the same pattern for different neighborhood sizes. From the plots, we also obtain that the frequencies of cooperators at the left bottom part stay at high level of proportion which are about or above 0.9, while those at the right top part are decreasing as the neighborhood size increases and most of them are less than 0.1. Furthermore we see that the average frequencies of cooperators will increase as the neighborhood size increases for the mutually beneficial game in last row with raw equilibrium frequency about 0.6. The frequencies on the last column are not so small, but there is a pattern of sharp decrease as  $N$  change from 8 to 24. Also we can see that the differences between the simulated frequencies and the original frequencies stay at the left bottom part, but become smaller for right top part as the neighborhood size increases. And this is consistent with the conclusion that coexistence can persist under localized competition and rapidly lost when dispersal and interaction enlarge in the C-S-R model promoted by Kerr, Riley, Feldman and Bohannan (10). Further more, as  $N$  becomes large and finally the interacting region extends to the whole lattice, the frequencies above the diagonal will converge to the original equilibrium frequencies which are almost zero.

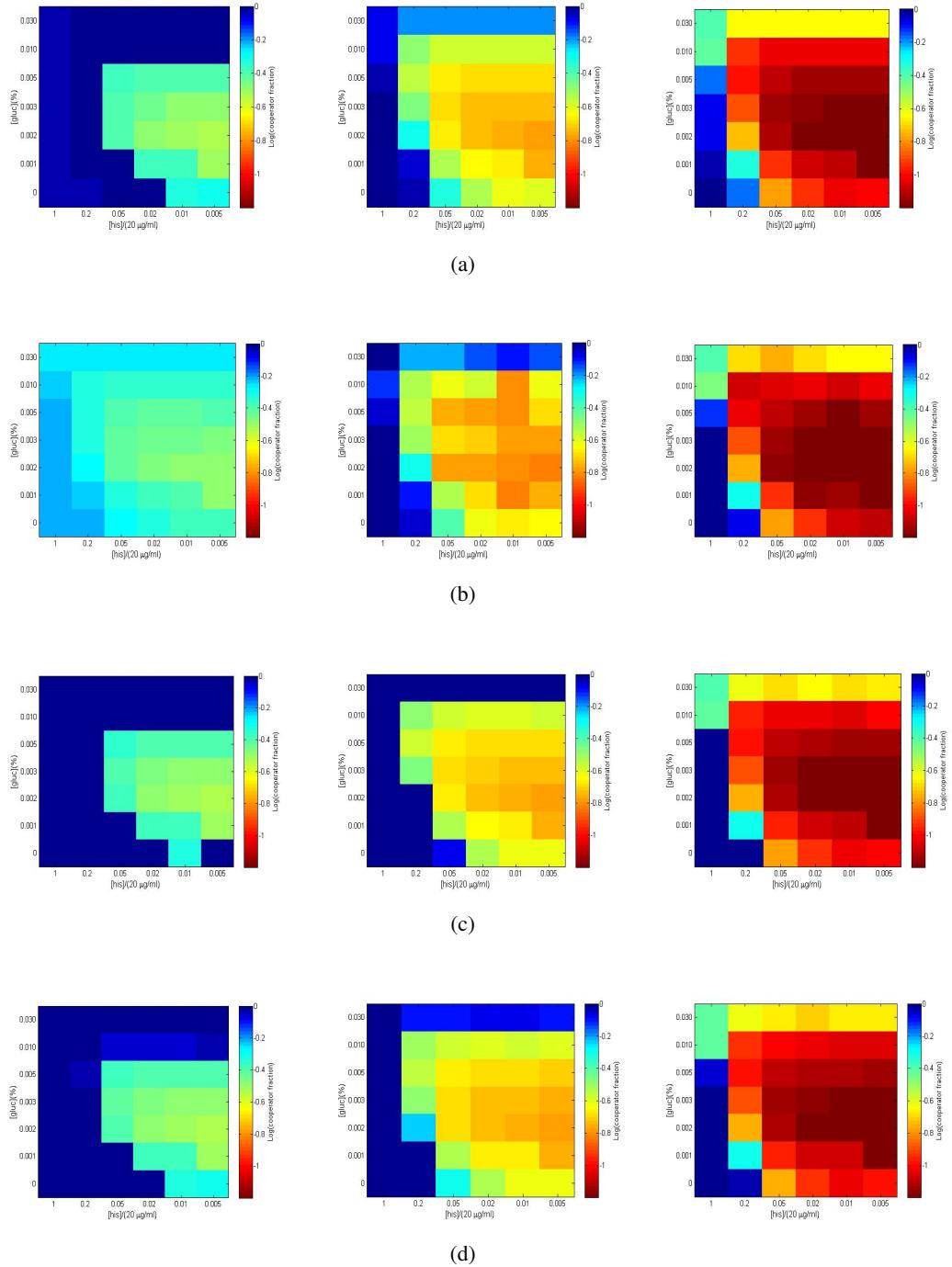


Figure 2.7 Plots for log frequencies of cooperators using deterministic rules in the numerical experiments based on four types of initialized lattices (3 columns from left to right correspond to  $N = 4, 8$  and  $24$ ). (a) Randomized initial lattice. (b) Initial lattice with a square of cooperators in a population of defectors. (c) Initial lattice with a square of defectors in a population of cooperators. (d) Initial lattice with left side cooperators and right side defectors.

### 2.4.2 Stochastic Rule

In stochastic update rule, the update probability is given in the form of  $p(x) = [1+\exp(-(P_C(N_x) - P_D(N_x))/\kappa)]^{-1}$  where  $x$  is site  $x$  on the lattice,  $N_x$  is the neighborhood of  $x$  and  $\kappa$  is a noise term that reflects uncertainties in assessing the payoffs which is usually set to be 0.01. The cooperator frequency plots for Prisoner's Dilemma and Snowdrift Game based on spatial evolutionary game dynamics using this stochastic update rule are displayed in Figure 2.8. They again tell us the approximate convergence of the average frequencies of cooperators. This shows that the average frequency of cooperators stays stable after  $M$  generations. The average frequencies of cooperators over certain generations after  $M$  generations for all cases are plotted and presented in Figure 2.8.

In Figure 2.8, different rows correspond to the simulation results of different initial lattices, and there is no obvious distinction among the four chosen cases. This indicates initial lattice has no effect on the evolution of cooperation in yeast game, which is the same as that in deterministic update. By comparing the average frequencies of certain generations after  $M$  generations using stochastic update rule with those using deterministic update rule, we can see that the stochastic updated "equilibrium" frequencies at the left bottom and right bottom parts, which are called deterministic-high-stochastic-low parts, are generally smaller than the deterministic updated ones under the same neighborhood size, while the situation is converse for the rest parts. There is also a pattern that the deterministic-high-stochastic-low parts are getting smaller as neighborhood size increases. Again the four different initial lattices give us almost the same "equilibrium" plots for the same neighborhood size which indicates the initial conditions generally don't affect the final stable states. The frequencies are getting larger as row number increases for each neighborhood size. We can further see that the average frequencies of cooperators will decrease as the neighborhood size increases for both Snowdrift games and Prisoner's Dilemma. Also the differences between the final static frequencies from simulation and the original frequencies are getting bigger at the left bottom part, but becoming smaller for right top part as the neighborhood size increases. And this is consistent with the conclusion that coexistence can persist under localized competition and rapidly lost when dispersal and interaction enlarge in the C-S-R model. Finally, as  $N$  becomes large and the interacting region for every individual extends to the whole lattice, the phenomena observed on the cooperator frequencies in the Prisoner's Dilemma games is the same

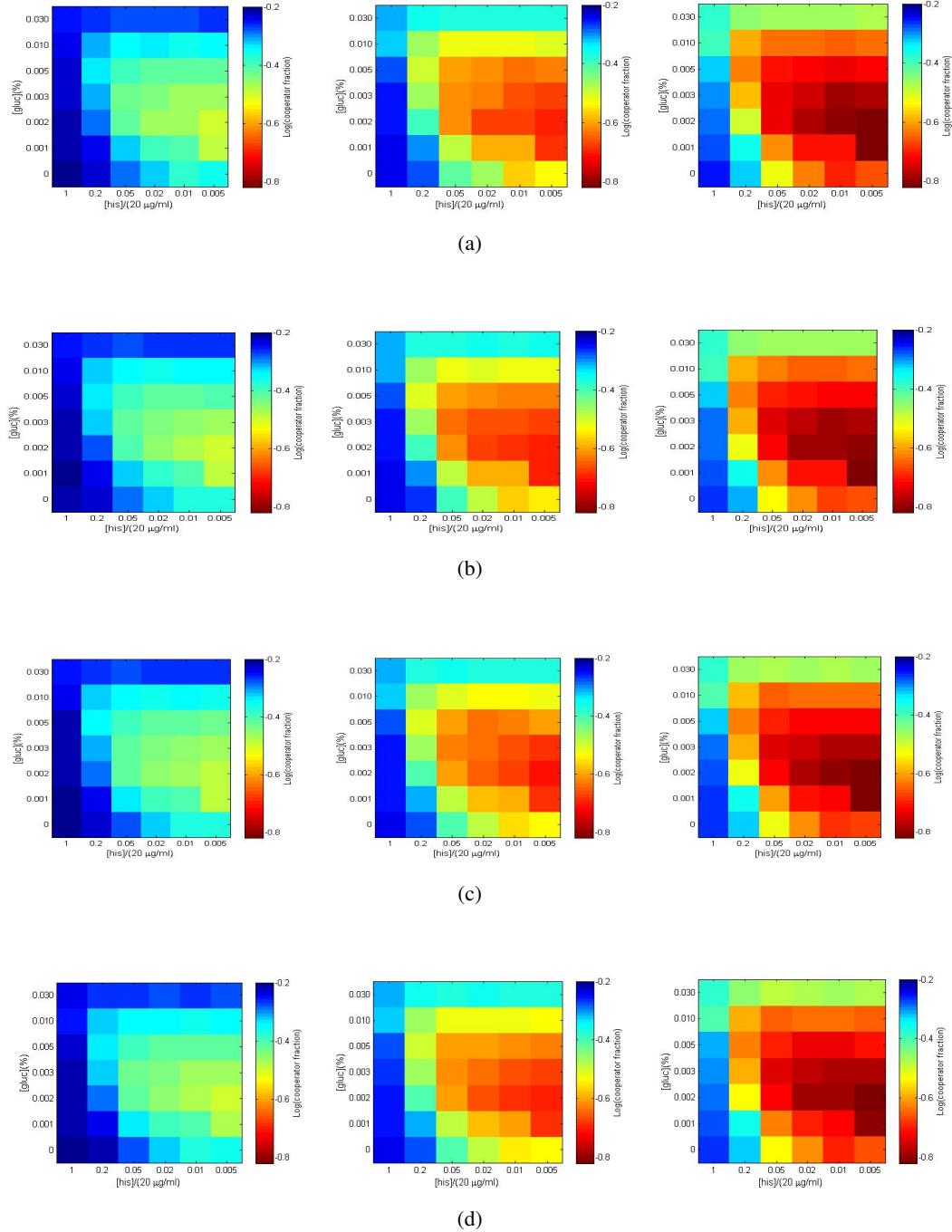


Figure 2.8 Plots for log frequencies of cooperators using stochastic rules in the numerical experiments based on four types of initialized lattices (3 columns from left to right correspond to  $N = 4, 8$  and  $24$ ). (a) Randomized initial lattice. (b) Initial lattice with a square of cooperators in a population of defectors. (c) Initial lattice with a square of defectors in a population of cooperators. (d) Initial lattice with left side cooperators and right side defectors.

as that observed by using deterministic rule, and they are all converging to the original equilibriums with corresponding frequencies that are almost zero.

## 2.5 General Mathematical Theory

For the 2D simulation in the yeast system, each grid on the lattice and its neighborhood region is denoted as a focal system. There are lots of focal systems on a 2D plane where the number of focal systems is the count of grids on the plane. There is a difference equation with discrete time and discrete space for each focal system, and they can be formulated as

$$\delta\langle f_i \rangle_\lambda(x, t)/\delta t = F\left(\langle f_i \rangle_\lambda(x, t), P_i(x, \langle f_i \rangle_\lambda(x, t)) - P(x, \langle f_i \rangle_\lambda(x, t))\right), \forall x,$$

where  $\lambda = \lambda(N)$ ,  $x$  is the lattice grid position,  $\langle f_i \rangle_\lambda(x, t) \in \{0, \frac{1}{M}, \frac{2}{M}, \dots, 1\}$  is the fraction of strategy or species  $i \in \{C, D\}$  in the system with  $M = N + 1$  grids,  $\delta t$  is the time difference between generations,  $\delta\langle f_i \rangle_\lambda(x, t)$  is fraction difference of strategy  $i$  between generations and  $F(\cdot, \cdot)$  characterizes the update probability of  $x$ . When  $F$  is of the form of indicator function, it corresponds to the deterministic update rule. When  $F$  is written as a continuous form and  $\delta t$  is infinitely small, the exponential form of the update probability could be approximately derived from the limit form of the difference equation, in which situation it corresponds to stochastic update rule. Specifically if  $F(x, y) = xy$  and the neighborhood extends to the whole lattice which is well-mixed population, then the difference equation boils down to limit form which is replicator equation as  $\delta t$  goes to 0. The difference equation is homogeneous for all focal grids. Also, the difference systems are not independent if the neighborhood regions of two focal grids are intersected, that means, the intersection part affects the systems which have the intersection part included. The effect will spread out to more systems and be weaker if the neighborhood size increases because there will be more systems involved.

To analyze the theory of the difference equation, we describe the system from the perspective of invasion circumstances based on at least two parameters—discrete time and discrete space in the perspective of deterministic update—defectors invading cooperators and cooperators invading defectors which are similar to single parameter case of the payoff matrix from Nowak book in (14), while in that of stochastic update it introduces many uncertainties which are put in a more complicated framework of future investigation.

### 2.5.1 Defectors Invading Cooperators

The following are directed edges from a directed graph for four cases of invaded defectors: (a) a single defector; (b) von Neumann neighborhood (Manhattan distance  $r = 1$ ) of 5 defectors; (c) square (Moore neighborhood) of 9 defectors; (d) von Neumann neighborhood (Manhattan distance  $r = 2$ ) of 13 defectors. For Prisoner's Dilemma or Snowdrift Game in yeast system, the payoffs  $P_C(f)$  and  $P_D(f)$  are both monotone increasing functions. The conditions for defectors-invading-cooperators under neighborhood size  $N = 4$  are the following:

- (i) (a)  $\rightarrow$  (b):  $\frac{1}{4}P_C(\frac{4}{5}) + \frac{3}{4}P_C(1) < P_D(\frac{4}{5})$ ,  $P_C(\frac{4}{5}) < P_D(\frac{4}{5})$ ;
- (ii) (b)  $\rightarrow$  (c):  $\frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(1) < P_D(\frac{3}{5}) < \frac{1}{4}P_C(\frac{4}{5}) + \frac{3}{4}P_C(1)$ ,  $\frac{2}{3}P_C(\frac{3}{5}) + \frac{1}{3}P_C(\frac{4}{5}) < \frac{1}{2}P_D(0) + \frac{1}{2}P_D(\frac{3}{5})$ ;
- (iii) (c)  $\rightarrow$  (d):  $\frac{3}{4}P_C(\frac{4}{5}) + \frac{1}{4}P_C(1) < P_D(\frac{1}{5})$ ,  $P_D(\frac{2}{5}) < \frac{1}{2}P_C(\frac{4}{5}) + \frac{1}{2}P_C(1)$ ,  $P_C(\frac{4}{5}) < \frac{2}{3}P_D(\frac{1}{5}) + \frac{1}{3}P_D(\frac{2}{5})$ ,  $P_C(\frac{4}{5}) < \frac{1}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5}) + \frac{1}{2}P_D(\frac{2}{5})$ ;
- (iv) (b)  $\rightarrow$  (d):  $\frac{1}{4}P_C(\frac{4}{5}) + \frac{3}{4}P_C(1) < P_D(\frac{3}{5})$ ,  $\frac{1}{3}P_C(\frac{4}{5}) + \frac{2}{3}P_C(1) < P_D(\frac{3}{5})$ ,  $\frac{2}{3}P_C(\frac{3}{5}) + \frac{1}{3}P_C(\frac{4}{5}) < \frac{1}{2}P_D(0) + \frac{1}{2}P_D(\frac{3}{5})$ ;
- (v) (d)  $\rightarrow$  (c):  $\frac{1}{2}P_D(0) + \frac{1}{2}P_D(\frac{3}{5}) < \frac{2}{3}P_C(\frac{3}{5}) + \frac{1}{3}P_C(\frac{4}{5})$ ,  $P_C(\frac{3}{5}) < \frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{2}{5})$ ,  $\frac{1}{2}P_D(\frac{2}{5}) + \frac{1}{2}P_D(\frac{3}{5}) < \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(1)$ ,  $P_D(\frac{3}{5}) < \frac{1}{4}P_C(\frac{4}{5}) + \frac{3}{4}P_C(1)$ ;
- (vi) (c)  $\rightarrow$  (b):  $P_C(\frac{4}{5}) < \frac{1}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5}) + \frac{1}{2}P_D(\frac{2}{5})$ ,  $\frac{2}{3}P_D(\frac{1}{5}) + \frac{1}{3}P_D(\frac{2}{5}) < P_C(\frac{4}{5})$ ,  $P_D(\frac{1}{5}) < \frac{3}{4}P_C(\frac{4}{5}) + \frac{1}{4}P_C(1)$ ,  $P_D(\frac{2}{5}) < \frac{1}{3}P_C(\frac{4}{5}) + \frac{2}{3}P_C(1)$ ;
- (vii) (b)  $\rightarrow$  (a):  $\frac{1}{2}P_D(0) + \frac{1}{2}P_D(\frac{3}{5}) < \frac{2}{3}P_C(\frac{3}{5}) + \frac{1}{3}P_C(\frac{4}{5})$ ,  $P_D(\frac{3}{5}) < \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(1)$ ,  $P_D(\frac{3}{5}) < \frac{1}{4}P_C(\frac{4}{5}) + \frac{3}{4}P_C(1)$ ;
- (viii) (d)  $\rightarrow$  (b):  $\frac{1}{2}P_D(0) + \frac{1}{2}P_D(\frac{3}{5}) < \frac{2}{3}P_C(\frac{3}{5}) + \frac{1}{3}P_C(\frac{4}{5})$ ,  $\frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{3}{5}) < P_C(\frac{3}{5})$ ,  $\frac{1}{2}P_D(\frac{2}{5}) + \frac{1}{2}P_D(\frac{3}{5}) < \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(1)$ ,  $P_D(\frac{3}{5}) < \frac{1}{4}P_C(\frac{4}{5}) + \frac{3}{4}P_C(1)$ ;
- (ix) (c)  $\rightarrow$  (a):  $\frac{1}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5}) + \frac{1}{2}P_D(\frac{2}{5}) < P_C(\frac{4}{5})$ ,  $\frac{2}{3}P_D(\frac{1}{5}) + \frac{1}{3}P_D(\frac{2}{5}) < P_C(\frac{4}{5})$ ,  $P_D(\frac{1}{5}) < \frac{3}{4}P_C(\frac{4}{5}) + \frac{1}{4}P_C(1)$ ,  $P_D(\frac{2}{5}) < \frac{1}{3}P_C(\frac{4}{5}) + \frac{2}{3}P_C(1)$ ;
- (x) (a)  $\rightarrow$  (a):  $P_D(\frac{4}{5}) < \frac{1}{4}P_C(\frac{4}{5}) + \frac{3}{4}P_C(1)$ ,  $P_C(\frac{4}{5}) < P_D(\frac{4}{5})$ ;
- (xi) (b)  $\rightarrow$  (b):  $P_D(\frac{3}{5}) < \frac{1}{4}P_C(\frac{4}{5}) + \frac{3}{4}P_C(1)$ ,  $P_D(\frac{3}{5}) < \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(1)$ ,  $\frac{2}{3}P_C(\frac{3}{5}) + \frac{1}{3}P_C(\frac{4}{5}) < \frac{1}{2}P_D(0) + \frac{1}{2}P_D(\frac{3}{5})$ ;
- (xii) (c)  $\rightarrow$  (c):  $P_D(\frac{1}{5}) < \frac{3}{4}P_C(\frac{4}{5}) + \frac{1}{4}P_C(1)$ ,  $P_D(\frac{2}{5}) < \frac{1}{2}P_C(\frac{4}{5}) + \frac{1}{2}P_C(1)$ ,  $P_C(\frac{4}{5}) < \frac{2}{3}P_D(\frac{1}{5}) + \frac{1}{3}P_D(\frac{2}{5})$ ;

$$\begin{aligned} \frac{1}{3}P_D\left(\frac{2}{5}\right), P_C\left(\frac{4}{5}\right) &< \frac{1}{4}P_D(0) + \frac{1}{4}P_D\left(\frac{1}{5}\right) + \frac{1}{2}P_D\left(\frac{2}{5}\right); \\ \text{(xiii) (d)} \rightarrow (d): \frac{2}{3}P_C\left(\frac{3}{5}\right) + \frac{1}{3}P_C\left(\frac{4}{5}\right) &< \frac{1}{2}P_D(0) + \frac{1}{2}P_D\left(\frac{3}{5}\right), P_C\left(\frac{3}{5}\right) < \frac{2}{3}P_D(0) + \frac{1}{3}P_D\left(\frac{2}{5}\right), \frac{1}{2}P_D\left(\frac{2}{5}\right) + \\ \frac{1}{2}P_D\left(\frac{3}{5}\right) &< \frac{1}{3}P_C\left(\frac{3}{5}\right) + \frac{2}{3}P_C(1), P_D\left(\frac{3}{5}\right) < \frac{1}{4}P_C\left(\frac{4}{5}\right) + \frac{3}{4}P_C(1). \end{aligned}$$

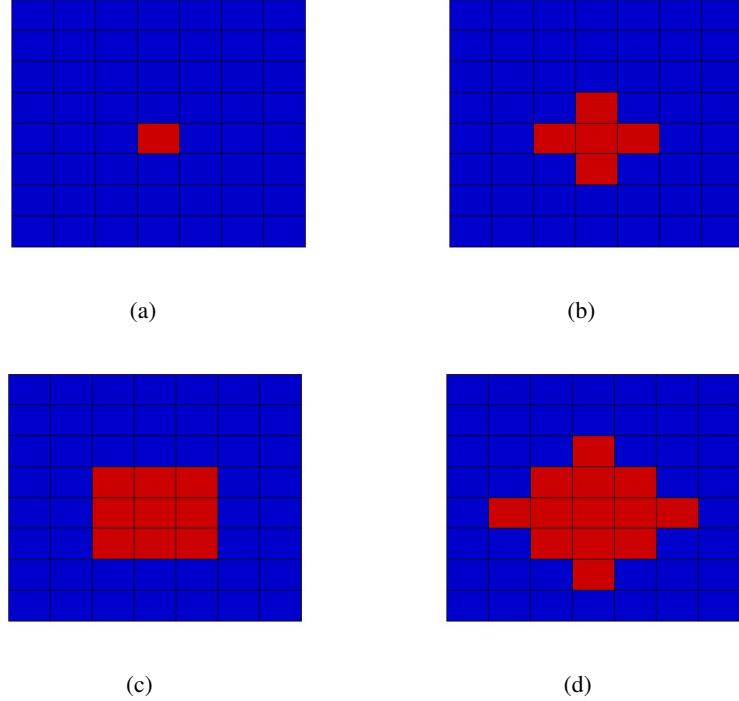


Figure 2.9 Invasion conditions for defectors.

The first four conditions are the inequalities correspond to the process of defectors' expansion. Conditions (v)-(ix) suggest the constraints while the defectors' region are shrinking, and conditions (x)-(xiii) tell us when the four cases stay still. It's interesting to consider how the 42 different combinations of parameters perform in the above 13 defectors-invading-cooperators conditions. We compute the values on the left hand side and right hand side of each inequality for all combinations of parameters, and obtain an impressive pattern for the True/False results of the inequalities which are attached in Tables 2.3 and 2.4.

For the defectors-increase conditions (i)-(iv), there is a general pattern that almost all 42 choices of parameters partially satisfy some inequalities and partially do not satisfy the rest inequalities, and this indicates some cases will change and other cases will stay or mutate to cases not listed above.

Furthermore, the parameter vectors correspond to the left top part all meet the conditions of (i), i.e., they tend to enlarge from a single cooperator to a von Neumann neighborhood of 5 cooperators in a population of defectors. But none of the parameters in the right and bottom parts satisfy condition (iv), i.e., there will never be the situation that a von Neumann neighborhood of 5 defectors expand to the von Neumann neighborhood of 13 defectors for the simulated models with parameters correspond to the right and bottom parts.

In addition, we find that the inequalities in defectors' shrinking conditions are mostly satisfied under all given sets of parameters. This suggests the risk of survival for defectors of 4 cases in an environment of surrounding cooperators because the cases with more defectors will be likely to switch to the cases with less defectors, which in some sense explains why coexistence exists in the simulation when the original game is Prisoner's Dilemma whose cooperator frequency is 0 and large portion of cooperators persist under localized competition.

Partially true and partially false results are also found in the inequalities of conditions (v)-(viii) under all parameter choices. To some extent, the results show us that it's hard for the 4 structures of defectors to keep themselves in a population of cooperators given any choice of parameters. These structures of defectors could keep portion of them sometimes, but the rest will change to structures which are more complicated than listed ones for most parameter combinations.

In summary, the above invasion description and analysis suggest there are no special patterns emerging if a single defector invades a world of cooperators for the given parameter choices. The resulting patterns are neither dynamical fractal that combines symmetry and chaos, nor "evolutionary kaleidoscope" which are found in linear 2D simulations. Actually they are locally cooperators preferred and mostly partially-keep-partially-mutate. And they also show us the deterministic unpredictability and asymmetry of beauty.

### 2.5.2 Cooperators Invading Defectors

We also provide four cases of invading cooperators: (a) a single cooperator; (b) von Neumann neighborhood (Manhattan distance  $r = 1$ ) of 5 cooperators; (c) square (Moore neighborhood) of 9 cooperators; (d) von Neumann neighborhood (Manhattan distance  $r = 2$ ) of 13 cooperators. The conditions for cooperators-invading-defectors under neighborhood size  $N = 4$  are the following:

- (i) (a)  $\rightarrow$  (b):  $\frac{3}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5}) < P_C(\frac{1}{5}), P_D(\frac{1}{5}) < P_C(\frac{1}{5});$
- (ii) (b)  $\rightarrow$  (c):  $\frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{2}{5}) < P_C(\frac{2}{5}) < \frac{3}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5}), \frac{1}{3}P_D(\frac{1}{5}) + \frac{2}{3}P_D(\frac{2}{5}) < \frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(1);$
- (iii) (c)  $\rightarrow$  (d):  $\frac{1}{4}P_D(0) + \frac{3}{4}P_D(\frac{1}{5}) < P_C(\frac{4}{5}), P_C(\frac{3}{5}) < \frac{1}{2}P_D(0) + \frac{1}{2}P_D(\frac{1}{5}), P_D(\frac{1}{5}) < \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(\frac{4}{5}), P_D(\frac{1}{5}) < \frac{1}{2}P_C(\frac{3}{5}) + \frac{1}{4}P_C(\frac{4}{5}) + \frac{1}{4}P_C(1);$
- (iv) (b)  $\rightarrow$  (d):  $\frac{3}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5}) < P_C(\frac{2}{5}), \frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{1}{5}) < P_C(\frac{2}{5}), \frac{1}{3}P_D(\frac{1}{5}) + \frac{2}{3}P_D(\frac{2}{5}) < \frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(1);$
- (v) (d)  $\rightarrow$  (c):  $\frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(1) < \frac{1}{3}P_D(\frac{1}{5}) + \frac{2}{3}P_D(\frac{2}{5}), P_D(\frac{2}{5}) < \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(1), \frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(\frac{3}{5}) < \frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{2}{5}), P_C(\frac{2}{5}) < \frac{3}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5});$
- (vi) (c)  $\rightarrow$  (b):  $P_D(\frac{1}{5}) < \frac{1}{2}P_C(\frac{3}{5}) + \frac{1}{4}P_C(\frac{4}{5}) + \frac{1}{4}P_C(1), \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(\frac{4}{5}) < P_D(\frac{1}{5}), P_C(\frac{4}{5}) < \frac{1}{4}P_D(0) + \frac{3}{4}P_D(\frac{1}{5}), P_C(\frac{3}{5}) < \frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{1}{5});$
- (vii) (b)  $\rightarrow$  (a):  $\frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(1) < \frac{1}{3}P_D(\frac{1}{5}) + \frac{2}{3}P_D(\frac{2}{5}), P_C(\frac{2}{5}) < \frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{2}{5}), P_C(\frac{2}{5}) < \frac{3}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5});$
- (viii) (d)  $\rightarrow$  (b):  $\frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(1) < \frac{1}{3}P_D(\frac{1}{5}) + \frac{2}{3}P_D(\frac{2}{5}), \frac{1}{3}P_C(\frac{2}{5}) + \frac{2}{3}P_C(1) < P_D(\frac{2}{5}), \frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(\frac{3}{5}) < \frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{2}{5}), P_C(\frac{2}{5}) < \frac{3}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5});$
- (ix) (c)  $\rightarrow$  (a):  $\frac{1}{2}P_C(\frac{3}{5}) + \frac{1}{4}P_C(\frac{4}{5}) + \frac{1}{4}P_C(1) < P_D(\frac{1}{5}), \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(\frac{4}{5}) < P_D(\frac{1}{5}), P_C(\frac{4}{5}) < \frac{1}{4}P_D(0) + \frac{3}{4}P_D(\frac{1}{5}), P_C(\frac{3}{5}) < \frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{1}{5});$
- (x) (a)  $\rightarrow$  (a):  $P_C(\frac{1}{5}) < \frac{3}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5}), P_D(\frac{1}{5}) < P_C(\frac{1}{5});$
- (xi) (b)  $\rightarrow$  (b):  $P_C(\frac{2}{5}) < \frac{3}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5}), P_C(\frac{2}{5}) < \frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{2}{5}), \frac{1}{3}P_D(\frac{1}{5}) + \frac{2}{3}P_D(\frac{2}{5}) < \frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(1);$
- (xii) (c)  $\rightarrow$  (c):  $P_C(\frac{4}{5}) < \frac{1}{4}P_D(0) + \frac{3}{4}P_D(\frac{1}{5}), P_C(\frac{3}{5}) < \frac{1}{2}P_D(0) + \frac{1}{2}P_D(\frac{1}{5}), P_D(\frac{1}{5}) < \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(\frac{4}{5}), P_D(\frac{1}{5}) < \frac{1}{2}P_C(\frac{3}{5}) + \frac{1}{4}P_C(\frac{4}{5}) + \frac{1}{4}P_C(1);$
- (xiii) (d)  $\rightarrow$  (d):  $\frac{1}{3}P_D(\frac{1}{5}) + \frac{2}{3}P_D(\frac{2}{5}) < \frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(1), P_D(\frac{2}{5}) < \frac{1}{3}P_C(\frac{3}{5}) + \frac{2}{3}P_C(1), \frac{1}{2}P_C(\frac{2}{5}) + \frac{1}{2}P_C(\frac{3}{5}) < \frac{2}{3}P_D(0) + \frac{1}{3}P_D(\frac{2}{5}), P_C(\frac{2}{5}) < \frac{3}{4}P_D(0) + \frac{1}{4}P_D(\frac{1}{5}).$

Similar as the defector-invading-cooperators conditions, we have the 13 conditions for cooperators-invading-defectors listed above. Conditions (i)-(iv) are the equivalent inequalities to cooperator clusters' growing. If conditions (v)-(ix) hold, then the defector clusters are taking over cooperators' sites. There

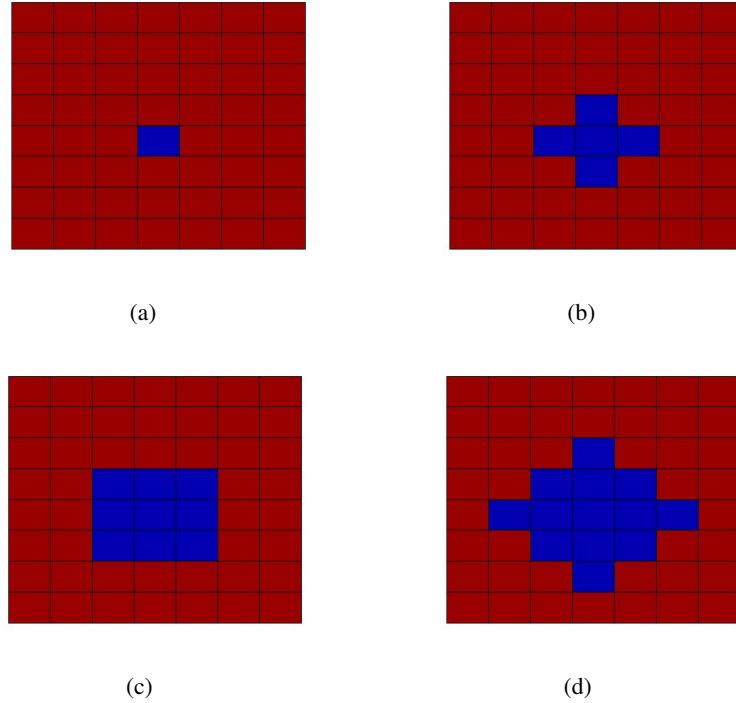


Figure 2.10 Invasion conditions for cooperators.

will be period one oscillator for all four cases when conditions (x)-(xiii) are satisfied. The performance of the 42 different combinations of parameters on the above 13 cooperators-invading-defectors conditions is again interesting. The computation on the left hand side and right hand side of each inequality for all parameters choices are Table 2.4, and the summary for the True/False results are as followed.

For the first four conditions, we encounter that most of the inequalities are true under all choices of parameters while a small proportion of them are false. This phenomena indicates that the single cooperator outperforms the neighbor defectors, and this single cooperator will keep growing to clusters with more cooperators, and a von Neumann neighborhood of 5 cooperators will certainly extend at each direction of its structure.

On the contrary, most of the inequalities in conditions (v)-(ix) don't hold for any given combination of parameters. We can expect that the clusters of cooperators are likely to expand more than just retaining themselves and have few motivation to shrink. In other words, they prefer to keep growing to structures of even more cooperators which are not listed here.

The True/False results of last four conditions (period one oscillator) are similar to those in the

situation of defectors invading cooperators, that is, part of the inequalities are true and the rest are false. More specifically, a single cooperator cannot retain itself and have the willing to expand; part of the cases (b), (c) and (d) can keep themselves, but the rest part cannot outperform defectors and their sites will be taken by defectors in the next generation.

To sum up, the invasion analysis tell us that the cooperators tend to grow if they are invading a population of defectors for any choice of parameter. The growing clusters of cooperators move around, expand, fragment and give birth to new clusters of cooperators. There is no special pattern emerging for a single cooperator invading a world of defectors. This also suggests that the successful dominance of cooperator under localized competition. Again the competition consists of deterministic unpredictability and beauty of asymmetry.

### 2.5.3 Generalization

The invasion analysis can be generalized in several aspects. The games can be studied with stochastic update rules and the invasion conditions will be unpredictable which are very complicated. Instead of asynchronous updating, we can investigate synchronous updating as well, and the basic structures will be symmetric. Synchronous updating introduces less random choice and therefore less stochasticity. It will in general display less variety in dynamical behaviors.

From the discussion for two situations of invasion with the neighborhood size  $N = 4$ , we can see that the 13 conditions are sufficient for the analysis of basic four structures. For general larger neighborhood sizes  $N = 8$  and  $24$ , the basic structures of cooperators or defectors are a little different, and thus there will be more conditions for the invasion of a single individual in a population with opposite strategy. It's fussy to list all conditions again for the situation of different neighborhood sizes. However we can anticipate that a single cooperator still tends to grow in a population of defectors and cooperator strategy could dominate in a localized competition but the dominance is not so strong as that with small neighborhood sizes.

## 2.6 Conclusion and Discussion

Evolutionary game theory has achieved tremendous success in modeling biological and social systems. However almost all of the successes have been focus on the linear games. The first applicable nonlinear game model is proposed by Gore et al. in 2009 and later the mathematical theory, especially the theoretical properties of this model is developed by us. We now make a further investigation of this nonlinear game model based on the perspective of  $2D$  simulation and the contributions of this paper can be listed as follows.

First of all, we have reviewed and generalized three types of linear games whose definitions are based on the comparative relationship of the elements in the so-called payoff matrix. The generalized nonlinear game are defined with nonlinear payoff functions instead of matrix-based linear functions. The Prisoner’s Dilemma and Snowdrift game for nonlinear games are conceptualized from the perspective of payoff functions’ curves and intercepts of payoff functions which are comparably same to those in matrix games, except the skewness and concavity of the curves. Parameters in payoff matrix as now replaced by environmental or internal parameters in nonlinear payoff functions. We also study the replicator dynamics of nonlinear games and would expect the nonlinear games can be widely use for a large group of systems.

Secondly, the combination of numerical simulation and evolutionary game theory has been proved to be useful to study the linear game models and how the spatial effects impact the evolution of co-operation. We extend the combined model to the field of nonlinear game models where rare scientist has considered before. We find that the spatial structure on a  $2D$  lattice tends to promote the evolution of cooperation in both Prisoner’s Dilemma and Snowdrift Game, which is a little different from those found in matrix games since Christoph shows that spatial effects increase the proportion of cooperation in Prisoner’s Dilemma while this doesn’t hold for Snowdrift Game in matrix games. We also compare the deterministic and stochastic update rules and conclude that their results are different from each other but tend to coincide as neighborhood size increases.

Thirdly, we propose the difference equations for the  $2D$  simulation that are implemented in Section 2.5. Then we analyze some basic properties of the difference equations for the systems of players. The invasion conditions are provided given four structures for two different situations with the neighborhood

size  $N = 4$ . We find some interesting patterns for the inequalities of the invasion conditions under different combinations of parameters from the experimental data Gore et.al collected and those patterns convince us the promotion of cooperation in the numerical experiments. We also make a generalization based on neighborhood sizes and update rules and concluding results are discussed.

Finally, this paper invites more interest from both mathematical and biological communities in evolutionary game theory and its applications in modeling. It's a further extension of the work about the game dynamic model for yeast development proposed by Huang and Wu (2012). Further development of our work can be fascinating, such as the rigorous theoretical properties for the 2D model on the lattice and so on.

### **Acknowledgements**

We'd like to thank Prof. Jeff Gore for his kindness by providing us the Matlab code for generating some of the graphical results in their paper and the helpful suggestions on how to proceed in our simulation.

Table 2.3 Defector invading cooperators (sign for inequalities: "+" means inequality holds while "–" means it doesn't hold.)

|       | (i) | (ii) | (iii) | (iv) | (v)  | (vi) | (vii) | (viii) | (ix) | (x) | (xi) | (xii) | (xiii) |
|-------|-----|------|-------|------|------|------|-------|--------|------|-----|------|-------|--------|
| (1,1) | ++  | ++-  | -+--  | ---  | +-++ | -+++ | +-+   | ++++   | ++++ | -+  | +-   | ++-   | --++   |
| (1,2) | ++  | +-   | ----  | ++-  | +-+- | -+++ | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (1,3) | ++  | +-   | ----  | ++-  | +-+- | -++- | +-    | ++-    | +++- | -+  | -    | -     | -      |
| (1,4) | ++  | +-   | -+--  | ++-  | +-+- | -++- | +-    | ++-    | +++- | -+  | -    | -     | -      |
| (1,5) | ++  | +-   | -+--  | ++-  | +-+- | -++- | +-    | ++-    | ++++ | -+  | -    | ++-   | -      |
| (1,6) | ++  | ++-  | -+--  | -    | +++  | -++  | +-    | +++    | ++++ | -+  | +    | -     | -++    |
| (1,7) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | ++-   | --++   |
| (2,1) | ++  | +-   | -+--  | ++-  | +-+- | -++  | +-    | ++-    | +++  | -+  | -    | -     | -      |
| (2,2) | ++  | +-   | -+--  | ++-  | +-+- | -++  | +-    | ++-    | +++  | -+  | -    | -     | -      |
| (2,3) | ++  | +-   | -+--  | ++-  | +-+- | -++  | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (2,4) | ++  | +-   | -+--  | ++-  | +-+- | -++  | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (2,5) | ++  | +-   | -+--  | ++-  | +-+- | -++  | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (2,6) | ++  | ++-  | -+--  | -    | +++  | -++  | +-    | +++    | ++++ | -+  | +    | -     | -++    |
| (2,7) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (3,1) | ++  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | -+  | ++-  | -     | -++    |
| (3,2) | ++  | -+-- | -+--  | ++-  | +-+- | -++  | +-    | ++-    | +++  | -+  | -    | -     | -      |
| (3,3) | ++  | -+-- | -+--  | ++-  | +-+- | -++  | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (3,4) | ++  | -+-- | -+--  | ++-  | +-+- | -++  | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (3,5) | ++  | -+-- | -+--  | ++-  | +-+- | -++  | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (3,6) | ++  | ++-  | -+--  | -    | +++  | -++  | +-    | +++    | ++++ | -+  | +    | -     | -++    |
| (3,7) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (4,1) | ++  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | -+  | ++-  | -     | -++    |
| (4,2) | ++  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | -+  | ++-  | -     | -++    |
| (4,3) | ++  | -+-- | -+--  | ++-  | +-+- | -++  | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (4,4) | ++  | -+-- | -+--  | ++-  | +-+- | -++  | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (4,5) | ++  | -+-- | -+--  | ++-  | +-+- | -++  | +-    | ++-    | ++++ | -+  | -    | -     | -      |
| (4,6) | ++  | ++-  | -+--  | -    | +++  | -++  | +-    | +++    | ++++ | -+  | ++-  | -     | -++    |
| (4,7) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (5,1) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (5,2) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (5,3) | ++  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | -+  | ++-  | -     | -++    |
| (5,4) | ++  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | -+  | ++-  | -     | -++    |
| (5,5) | ++  | ++-  | -+--  | -    | +++  | -++  | +-    | +++    | ++++ | -+  | ++-  | -     | -++    |
| (5,6) | ++  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | -+  | ++-  | -     | -++    |
| (5,7) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (6,1) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (6,2) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (6,3) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (6,4) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (6,5) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (6,6) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |
| (6,7) | -+  | -+-- | -+--  | -    | +++  | -++  | ++    | +++    | ++++ | ++  | ++-  | -     | -++    |

Table 2.4 Cooperator invading defectors (sign for inequalities: "+" means inequality holds while "−" means it doesn't hold.

|       | (i) | (ii) | (iii)  | (iv) | (v)    | (vi)   | (vii) | (viii)  | (ix)  | (x)   | (xi)   | (xii)   | (xiii) |
|-------|-----|------|--------|------|--------|--------|-------|---------|-------|-------|--------|---------|--------|
| (1,1) | ++  | +-+  | + - ++ | ++ + | - + -- | + - -- | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (1,2) | ++  | +- - | + - ++ | ++ - | + - -- | + - -- | + - - | + + - - | - - - | - - - | - - ++ | - - - - |        |
| (1,3) | ++  | +- - | + - ++ | ++ - | + - -- | + - -  | + - - | + + - - | - - - | - - - | - - ++ | - - - - |        |
| (1,4) | ++  | +- - | + - ++ | ++ - | + + -- | + - -  | + - - | + + - - | - - - | - - - | - - ++ | - + - - |        |
| (1,5) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (1,6) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (1,7) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (2,1) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (2,2) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - - | - - ++ | ++ - -  |        |
| (2,3) | ++  | +- - | + - ++ | ++ - | + - -- | + - -  | + - - | + + - - | - - - | - - - | - - ++ | - - - - |        |
| (2,4) | ++  | +- - | + - ++ | ++ - | + + -- | + - -  | + - - | + + - - | - - - | - - - | - - ++ | - + - - |        |
| (2,5) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (2,6) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (2,7) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (3,1) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (3,2) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (3,3) | ++  | +- - | + - ++ | ++ - | + + -- | + - -  | + - - | + + - - | - - - | - - - | - - ++ | - + - - |        |
| (3,4) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - + - - | - - - | - - + | - - ++ | ++ - -  |        |
| (3,5) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (3,6) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (3,7) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (4,1) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (4,2) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (4,3) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (4,4) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (4,5) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (4,6) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (4,7) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (5,1) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (5,2) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (5,3) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (5,4) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (5,5) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (5,6) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (5,7) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - ++ | ++ - -  |        |
| (6,1) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - +   | - - +  | - - ++  | ++ - - |
| (6,2) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - +   | - - +  | - - ++  | ++ - - |
| (6,3) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - +   | - - +  | - - ++  | ++ - - |
| (6,4) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - +  | - - ++  | ++ - - |
| (6,5) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - +  | - - ++  | ++ - - |
| (6,6) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - +  | - - ++  | ++ - - |
| (6,7) | ++  | +- + | + - ++ | ++ + | - + -- | + - -  | - - - | - - -   | - - - | - - + | - - +  | - - ++  | ++ - - |

## CHAPTER 3. ESS IN NONLINEAR GAME AND THE ASYMPTOTIC BEHAVIOR OF SOLUTION TO ADAPTED REPLICATOR DIFFUSION EQUATION

A paper to be submitted

Min Wang, Zhijun Wu

### **Abstract**

In this paper we introduce the generalized replicator dynamics in game theory with nonlinear payoff functions of  $n$  strategies for all species in a population. Thus the necessary and sufficient conditions for the payoff functions under the evolutionarily stable state and regular evolutionarily stable state are investigated in the framework of evolutionary game theory. Furthermore, we extend the generalized replicator dynamics to adapted replicator diffusion equation by adding an diffusion term, and propose the theorem concerning the long time behavior of the solution to the replicator diffusion equation under appropriate conditions on the payoff functions. Simulation studies of two examples of replicator diffusion equation are followed to verify the asymptotic behavior of the replicator diffusion equation.

### **3.1 Introduction**

Von Neumann and Morgenstern (1944) defined the classical game theory by introducing the axioms for the concept of the individual rational player in the special case of zero-sum game in (39). But the player who made consistent decisions did not necessarily assume the other players also acted rationally. In 1951, John Nash set up the modern game theory by defining rational players in a degree such that they could coordinate their strategies where a mixed strategy Nash equilibrium prevailed in (34). Later on, Maynard Smith and Price (1973) formulated the evolving populations in a way where competition of several species could be analyzed as strategies and some mathematical criteria were used to predict

the prevalence of the competing strategies in (32). Then Taylor and other scientists focused more on the dynamics of the strategy which depended on the relative abundances or frequencies of all species in the population (30; 38; 41). This is the origination of the evolutionary game theory.

A fundamental goal of evolutionary game theory is to understand how the individual organisms in biological system or social system interact with each other, and evolve over time with the environment determine the spatial distribution of populations and the structure of communities. It considers a population of players interacting in a game with fixed strategies and payoff of each individual is interpreted as fitness, and success in the game is translated into reproductive success. The reproduce process is described by the evolutionary game dynamics and the paper of Taylor and Jonker (1978) used a differential equation for game dynamics which is the pioneer work of this area. The differential equation is called replicator dynamic equation and the payoff of a strategy in the equation is an estimate of the growth rate of players choosing this strategy. If the payoff functions of all strategies are linear combinations of the frequencies of all species, then the game is called matrix game or linear game. Furthermore, Zeeman (1980), Schuster , Hofbauer, and Sigmund (1979), and so on investigated on the properties of replicator equation in such game and there were collective research results on this equation. However few have considered the situation that at least one of the payoff functions are actually nonlinear functions or more general functions with respect to the frequencies of all species with the corresponding game called nonlinear game. We will firstly focus on the theory of this general replicator dynamics for nonlinear games in the following section.

A basic task of evolutionary game theory is to explore the stability of game dynamics. Nash equilibrium is a traditional concept in the stability theory of game which predicts the outcome of a strategic interaction. If there is a set of strategies with the property that no player can benefit by changing her strategy while the other players keep their strategies unchanged, then that set of strategies constitute the Nash equilibrium. Evolutionarily stable strategy (ESS) is another fundamental concept of evolutionary game theory, and it is introduced by Maynard Smith and Price in terms of two-component populations in a 1973 nature paper. ESS is an equilibrium refinement of Nash equilibrium. An ESS is a strategy such that, if most of the members of a population adopt it, there is no mutant strategy that would give higher reproductive fitness, that is, it is resistant to invasion by mutant strategies. Given the definition of ESS, there are lots of literature that analyze the conditions of the payoff matrix under which the ESS

exist in a linear game for different type of matrices, for example, Haigh(1975), Bishop and Cannings (1976), (1978), Taylor and Jonker (1978), Abakuks (1980) and so on (19; 20; 21; 28; 38). But there is rare result on the necessary and sufficient conditions for ESS with respect to the differentiation of payoff functions in a replicator dynamics of nonlinear games. We thus derive the equivalent conditions of ESS in terms of the derivatives of payoff functions and the equilibrium strategy using the Taylor expansion in this paper.

In recent decades, empirical evidence suggests that the spatial structure of environments in a population can influence interactions of species and the composition of communities. The natural way to understand spatial effects is to use mathematical models. There are several types of spatial population models, for example, cellular automaton, interacting particle systems, the idea free distribution and dispersal models based on integral kernels. Each type is based on some corresponding hypothesis about the scale and structure of the spatial environments and the way the species disperse through it. Diffusion term in the general reaction-diffusion equations also provides a good perspective to incorporate the spatial effects into a differential equation model. It translates the local assumptions or data about the movement strategy and reproductivity of individuals into global considerations of the competition and cooperation of population species. For the replicator differential equation of nonlinear games we mentioned above, we add the diffusion term and adapt the replicator dynamics to replicator diffusion equation with appropriate initial and boundary conditions. The goodness of this adaptation is that it spatially includes quantities such as dispersal rates, carrying capacities as parameters in the partial differential equation. Thus we can study questions about the way that variation in parameters especially diffusion rates influences the population dynamics.

Stability is a fundamental goal for the evolutionary game theory. That is, we care about the convergence of the solution of the game dynamics. Thus the long time behavior of solution of the adapted replicator diffusion equation is an interesting problem to us. The adapted replicator diffusion equation belongs to the class of general reaction diffusion equation and there are already some theoretical results for the long time behavior of reaction diffusion equation. Friedman (1964) considered a special class of reaction diffusion equations called the first and second initial boundary value problem and proved the theorems for the convergence of solutions of these two type of parabolic equations. Ewer and Peletier (1975), Chafee (1975), Mimura (1979), Redheffer, Redlinger and Walter (1988) and others examined

the asymptotic behavior of special cases of different parabolic equations with reasonable assumptions which guaranteed the convergence of the solutions and obtained the asymptotic stable conclusions of the equilibrium solutions (22; 23; 33; 37). However the adapted replicator diffusion equation we established does not satisfy the specific assumptions of the above results. For the linear game, Hadeler (1981) adapted it to corresponding replicator diffusion equation and proposed a theorem that the solution of his adapted equation converges uniformly in a bounded region towards a spatially homogeneous function which is an exponentially stable stationary state of the ordinary differential equation in (27). We will make a generalization of Hadeler's result. For the adapted replicator diffusion equation, we show that this resulting parabolic system with zero flux boundary conditions has the solution that converges to a spatially homogeneous stationary solution of the original replicator dynamics under appropriate conditions.

The rest of this paper is organized as follows. In section 3.2 we introduce the ESS of nonlinear replicator dynamics and derive the necessary and sufficient conditions of payoff functions for ESS. In section 3.3 some lemmas are set up and the convergence of the solution of the adapted replicator diffusion equation is proved. Some examples and numerical simulation results are provided in section 3.4. Finally we sum up with the remarks and conclusions.

### 3.2 ESS in Nonlinear Games

In 1973 Maynard Smith and Price introduced the concept of evolutionarily stable strategy (ESS) to describe a stable state of the game. Taylor and Jonker (1978) considered a general class of games which usually include the linear game and provided a general definition of ESS. They defined  $F(i|s)$  be the fitness of strategy  $i$  in state  $s$  where  $s = (s_1, \dots, s_n)$  is the state vector of the population and  $F(q|p) = \sum q_i F(i|p)$ ,  $s \in S = \{s \in R^n : \sum_{i=1}^n s_i = 1, s_i \geq 0, i = 1, \dots, n\}$ . Then a state  $p$  is called an *ESS* if for every state  $q \neq p$ , if we let  $\bar{p} = (1 - \epsilon)p + \epsilon q$ , then  $F(q|\bar{p}) < F(p|\bar{p})$  for sufficiently small  $\epsilon > 0$ . Define the matrix  $A$  by the formula

$$a_{ij} = \frac{\partial}{\partial p_j} F(i|p).$$

Then an equilibrium point  $p$  is a *regular ESS* if  $F(i|p) < F(p|p)$  whenever  $i \notin \text{supp}(p)$ , and  $x^T A x < 0$  whenever  $\text{supp}(x) \subset \text{supp}(p)$ ,  $x \neq 0$ , and  $\sum x_i = 0$ .

For linear game with payoff matrix  $U$ , Hofbauer and Sigmund (1998) denote the payoff for a  $p$ -strategist against a  $q$ -strategist to be

$$p^T \cdot Uq = \sum u_{ij}p_i q_j.$$

Thus the strategy  $\hat{p}$  is said to be evolutionarily stable if for all  $p$  with  $p \neq \hat{p}$ , the inequality

$$p^T \cdot U(\epsilon p + (1 - \epsilon)\hat{p}) < \hat{p}^T \cdot U(\epsilon p + (1 - \epsilon)\hat{p}) \quad (3.1)$$

holds for all sufficient small  $\epsilon > 0$ . The above inequality is a special case of the definition of ESS in Taylor and Jonker. Based on the form of linearity on  $\epsilon$ , the above inequality is equivalent to the following two conditions:

(a). *equilibrium condition*

$$p^T \cdot U\hat{p} \leq \hat{p}^T \cdot U\hat{p}.$$

(b). *stability condition*

$$p \neq \hat{p}, p^T \cdot U\hat{p} = \hat{p}^T \cdot U\hat{p} \implies p^T \cdot Up < \hat{p}^T \cdot Up.$$

Actually Hofbauer mentioned that the ESS is also equivalent to

$$\hat{p}^T \cdot Uq > q^T \cdot Uq$$

for all  $q \neq \hat{p}$  in some neighborhood of  $\hat{p}$ .

However for the nonlinear replicator dynamics, it's not easy to find the appropriate condition to guarantee the existence of ESS. In this section, we want to derive the necessary and sufficient conditions for ESS of the nonlinear replicator dynamics:

$$\frac{dp_i}{dt} = p_i(e_i^T \cdot \pi(p) - p^T \cdot \pi(p)), \quad i = 1, \dots, n, \quad (3.2)$$

where  $\pi(p) = (\pi_1(p), \dots, \pi_n(p))^T$  is a vector of bounded nonlinear payoff functions. Assume the payoff vector  $\pi(\cdot)$  is  $C^\infty$  at a small neighborhood of  $p$ . Rewrite the inequality for ESS  $p$  defined by Taylor and Jonker, we obtain

$$p^T \cdot \pi((1 - \epsilon)p + \epsilon q) > q^T \cdot \pi((1 - \epsilon)p + \epsilon q), \quad (3.3)$$

where  $\epsilon > 0$  is sufficiently small. Take the Taylor expansion for the term  $\pi((1 - \epsilon)p + \epsilon q)$  at point  $p$ , we have

$$\begin{aligned}\pi((1 - \epsilon)p + \epsilon q) &= \pi(p + \epsilon(q - p)) \\ &= a_0 + a_1\epsilon + \frac{1}{2}a_2\epsilon^2 + \cdots + \frac{1}{n!}a_n\epsilon^n + \cdots\end{aligned}$$

where  $a_0 = \pi(p)$ ,  $a_1 = \nabla\pi(p) \cdot (q - p)$ ,  $a_2 = \text{diag}((q - p)^T, \dots, (q - p)^T) \text{diag}(\nabla^2\pi_1(p), \dots, \nabla^2\pi_n(p))$   
 $1_{n \times 1} \otimes (q - p), \dots$

Thus

$$p^T \cdot \pi((1 - \epsilon)p + \epsilon q) = \sum_{i=0}^{\infty} \frac{1}{i!} p^T a_i \epsilon^i \quad (3.4)$$

and

$$q^T \cdot \pi((1 - \epsilon)p + \epsilon q) = \sum_{i=0}^{\infty} \frac{1}{i!} q^T a_i \epsilon^i. \quad (3.5)$$

In algebra, we know that if two infinite series  $\sum_{n=0}^{\infty} c_n \epsilon^n$  and  $\sum_{n=0}^{\infty} d_n \epsilon^n$  satisfy

$$\sum_{n=0}^{\infty} c_n \epsilon^n > \sum_{n=0}^{\infty} d_n \epsilon^n$$

for sufficient small  $\epsilon > 0$ , then the inequality implies that

- (1a)  $c_0 \geq d_0$ ;
- (2a) If  $c_0 = d_0$ , then  $c_1 \geq d_1$ ;
- (3a) If  $c_0 = d_0, c_1 = d_1$ , then  $c_2 \geq d_2$ ;
- $\vdots$

Based on the above result, we can derive that the ESS of the nonlinear replicator dynamics is equivalent to the following set of conditions:

- (1b)  $p^T a_0 \geq q^T a_0$ ;
- (2b) If  $p^T a_0 = q^T a_0$ , then  $p^T a_1 \geq q^T a_1$ ;
- (3b) If  $p^T a_0 = q^T a_0, p^T a_1 = q^T a_1$ , then  $p^T a_2 \geq q^T a_2$ ;
- $\vdots$

That is,

- (1c)  $p^T \pi(p) \geq q^T \pi(p)$ ;
- (2c) If  $p^T \pi(p) = q^T \pi(p)$ , then  $p^T \cdot \nabla\pi(p)(q - p) \geq q^T \cdot \nabla\pi(p)(q - p)$ ;
- (3c) If  $p^T \pi(p) = q^T \pi(p), p^T \cdot \nabla\pi(p)(q - p) = q^T \cdot \nabla\pi(p)(q - p)$ , then  $p^T \cdot \text{diag}((q - p)^T, \dots, (q - p)^T) \geq q^T \cdot \text{diag}((q - p)^T, \dots, (q - p)^T)$

$$\begin{aligned}
& p)^T) \text{diag}(\nabla^2 \pi_1(p), \dots, \nabla^2 \pi_n(p)) \mathbf{1}_{n \times 1} \otimes (q - p) \geq q^T \cdot \text{diag}((q - p)^T, \dots, (q - p)^T) \text{diag}(\nabla^2 \pi_1(p), \dots, \\
& \nabla^2 \pi_n(p)) \mathbf{1}_{n \times 1} \otimes (q - p); \\
& \vdots
\end{aligned}$$

Define the set of stationary points satisfy the set of conditions including the first  $(k - 1)$  ones and the  $k$ th one with inequality strictly holds to be  $\mathcal{E}_k$ , and the stationary points satisfy all the above set of infinite conditions to be  $\mathcal{E}_\infty$ . Denote the set of ESS to be  $\mathcal{E}$ . Based on the above derivation, we propose the following theorem which is necessary and sufficient for the existence of ESS of nonlinear game.

**Theorem 3.2.1.** *Consider the general replicator dynamics (3.2) of nonlinear game for  $n$  species with  $C^\infty$  payoff functions in a population. Then  $p$  is an ESS of the game iff  $p$  is a stationary state of (3.2) and  $p \in \mathcal{E}_\infty$ , that is, for any  $q \in S$  and  $q \neq p$ ,  $p$  satisfy all the conditions (1c), (2c), (3c), ... .*

It's hard to test all the conditions to verify whether a stationary state  $p$  is an ESS or not. Thus it's necessary to consider subsets of the conditions which are stronger than those in theorem 3.2.1. And we can see that the sequence of  $\mathcal{E}_k$  are monotonically nondecreasing as  $k \rightarrow \infty$ . However when the payoff functions are  $C^N$  for some integer  $N > 0$ , the conditions after  $(N + 1)$ th one cannot be verified. Therefore we have the following theorem

**Theorem 3.2.2.** *For the general replicator dynamics of nonlinear game to  $n$  species with  $C^\infty$  payoff functions, the stationary point sets  $\mathcal{E}_k$  are ESS and the sequence of them satisfy the two conditions:  $\mathcal{E}_k \subset \mathcal{E}_{k+1}$ , and  $\mathcal{E}_k \subset \mathcal{E}_\infty = \mathcal{E}$  for each  $k \geq 1$ . If the payoff functions are  $C^N$  for some integer  $N > 0$ , then  $\mathcal{E}_N \subset \mathcal{E}_k = \mathcal{E}_{N+1}$  for  $k \geq N + 1$ , and  $\mathcal{E}_1 \subset \mathcal{E}_2 \subset \dots \subset \mathcal{E}_{N-1} \subset \mathcal{E}_N \subset \mathcal{E}_{N+1} \subset \mathcal{E}$ .*

We can see that when  $\pi(p)$  is a matrix-times-vector form in terms of variable  $p$ , then the first two conditions are the same as the equilibrium condition and stability condition Hofbauer and Sigmund propose, that is,  $\mathcal{E}_k = \mathcal{E}_2$  for  $k \geq 2$ . However when  $\pi(p)$  is not of that form, that is,  $\pi(p)$  is nonlinear with respect to  $p$ , there are infinite conditions for the nonlinear payoff functions in theorem 3.2.1, and that's why Taylor and Jonker consider the stronger form of first two inequalities and introduce the concept of regular ESS which is also stronger than ESS. That is, we have the following theorem.

**Theorem 3.2.3.** *For the same replicator dynamics (3.2) of nonlinear game,  $p$  is a regular ESS of the game only if  $p$  is a stationary state of (3.2), and  $p \in \mathcal{E}_2$ , that is, for any  $q \in S$  and  $q \neq p$ , the*

two conditions hold: (1)  $p^T \pi(p) \geq q^T \pi(p)$ ; (2) If  $p^T \pi(p) = q^T \pi(p)$ , then  $p^T \cdot \nabla \pi(p)(q - p) > q^T \cdot \nabla \pi(p)(q - p)$ .

*Proof.*  $p$  is a regular ESS, then it is also an ESS. Thus  $p$  is a stationary state of (3.2). By definition of regular ESS, we have  $p^T \pi(p) > e_i^T \pi(p)$  where  $i \notin \text{supp}(p)$ . Then for any  $q \in S$  and  $q \neq p$ , we obtain

$$\begin{aligned} q^T \pi(p) &= \sum_{i \notin \text{supp}(p)} q_i \cdot e_i^T \pi(p) + \sum_{i \in \text{supp}(p)} q_i \cdot e_i^T \pi(p) \\ &\leq \sum_{i \notin \text{supp}(p)} q_i \cdot p^T \pi(p) + \sum_{i \in \text{supp}(p)} q_i \cdot e_i^T \pi(p) \\ &= \sum_{i \notin \text{supp}(p)} q_i \cdot p^T \pi(p) + \sum_{i \in \text{supp}(p)} q_i \cdot p^T \pi(p) \\ &= p^T \pi(p) \end{aligned}$$

where the equality holds when  $\text{supp}(q) \subset \text{supp}(p)$ . If  $p^T \pi(p) = q^T \pi(p)$ , then let  $x = q - p$ , we obtain the second condition.  $\square$

### 3.3 Long Time Behavior of Solution to the Adapted Replicator Diffusion Equation

We consider the replicator dynamics model of population for  $n$  species where  $p_i$  is the frequency of the  $i$ -th species, and the payoff function for  $i$ -th species is  $\pi_i(p)$  which is locally Lyapunov continuous. Then the differential equations is

$$\frac{dp_i}{dt} = p_i(e_i^T \cdot \pi(p) - p^T \cdot \pi(p)) \quad (3.6)$$

for  $i = 1, \dots, n$ . Consider the population distributed in a bounded domain  $\Omega$  of  $R^N$  with smooth boundary and incorporate migration and reproductive interaction of neighboring individuals by a diffusion term with the same diffusion rates

$$\frac{\partial p_i}{\partial t} = p_i(e_i^T \cdot \pi(p) - p^T \cdot \pi(p)) + \Delta p_i \quad (3.7)$$

for  $i = 1, \dots, n$ . That is,

$$\frac{\partial p}{\partial t} = P\pi(p) - p^T \pi(p) \cdot p + \Delta p \quad (3.8)$$

where  $P = \text{diag}(p_1, \dots, p_n)$ . The diffusion rate must be the same for all species, therefore we can normalized it to one by rescaling the space variables. Also we assume no flux condition on the boundary

$\partial\Omega$  and  $\mathbf{n}$  is the outer normal vector

$$\frac{\partial p}{\partial \mathbf{n}} \Big|_{\partial\Omega} = 0. \quad (3.9)$$

This corresponds to a situation where there is some population reservoir with reflecting walls. Hadeler (1981) considered the linear payoff function  $\pi_i(p) = \sum_{j=1}^n f_{ij}p_j$  where  $F = (f_{ij})$  is the symmetric payoff matrix with positive elements, thus the replicator dynamics is

$$\frac{dp}{dt} = PFp - p^T Fp \cdot p. \quad (3.10)$$

After adding the diffusion term, the differential equations becomes

$$\frac{\partial p}{\partial t} = PFp - p^T Fp \cdot p + \Delta p. \quad (3.11)$$

If all types of species are present, i.e.,  $p_i(x, t) \neq 0$  for  $i = 1, \dots, n$ , then this stationary solution is called the *polymorphism*. For the adapted linear replicator diffusion equation (3.11), Hadeler investigated the long time behavior of the solution and proposed the following theorem.

**Theorem 3.3.1.** *Let the ordinary differential equation (3.10) have an exponentially stable stationary state with all components positive (stable polymorphism). Then for every initial function  $p(\cdot, 0)$ , the solution of the replicator diffusion equation (3.11) converges uniformly in  $\Omega$  towards a spatially homogeneous (i.e., constant) function, which is a stationary state of the ordinary differential equation.*

The main tool used in the proof is the establishment of the spatial Lyapunov function and the comparison theorem for heat equations. We now state some preliminaries about these.

**Lemma 3.3.2.** *(Comparison principle) For the heat equation*

$$u_t - \Delta u = f \quad (3.12)$$

*in a region  $\Omega$  with initial condition  $u(x, 0) = g$  and Neumann boundary condition  $\frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial\Omega} = h$ . If we have two solutions  $u_1$  and  $u_2$  with data satisfying  $f_1 \geq f_2$ ,  $g_1 \geq g_2$  and  $h_1 \geq h_2$ , then  $u_1 \geq u_2$  everywhere.*

This lemma can be proved via the maximum principle. The following lemma will state the classical result about the stabilization of the solution of the Neumann problem for the heat equation with zero source .

**Lemma 3.3.3.** *Consider the heat equation*

$$\begin{aligned} u_t - \Delta u &= 0, \quad (x, t) \in \Omega \times R^+; \\ u(x, 0) &= g, \quad x \in \Omega; \\ \frac{\partial u}{\partial \mathbf{n}} &= 0, \quad (x, t) \in \partial\Omega \times R^+. \end{aligned} \tag{3.13}$$

If  $u(x, t)$  is the solution of the equation and the initial function  $g$  satisfies  $\int_{\Omega} |g| dx \leq M$  for some  $M > 0$ , then  $u(x, t) \rightarrow \bar{u}$  as  $t \rightarrow \infty$  where  $\bar{u} = \int_{\Omega} g dx / |\Omega|$ .

The above lemma can be proved by the method of separation of variables and fourier series. Next we are interested in the replicator diffusion equation (3.7) or (3.8) where the replicator term depends generalized nonlinear functions. Consider the initial condition

$$p(x, 0) = p_0(x) \tag{3.14}$$

where  $p_0(x) \in C^{2+\gamma}(\bar{\Omega})$  for some  $\gamma > 0$ . Redheffer et.al. (1983) provided the theorem about the existence and uniqueness of the solution of the general initial-boundary value problem in (37). This theorem is the following lemma.

**Lemma 3.3.4.** *Let  $\Omega$  be a bounded open connected subset of  $R^n$  with orientable boundary  $\partial\Omega$  of class  $C^{2+\gamma}$ . Assume that  $D \subset R^n$  is open and  $f : D \rightarrow R^n$  is locally Lipschitz continuous, also the matrix  $A(x) = (a_{ij}(x))$  is of class  $C^{1+\gamma}(\bar{\Omega})$  and  $A \geq \alpha I_m$  with  $\alpha > 0$ . Furthermore,  $\bar{u}(x) \in C^{2+\gamma}(\bar{\Omega})$ ,  $\partial\bar{u}/\partial n = 0$  on  $\partial\Omega$  and  $\bar{u}(x) \in D$  for  $x \in \bar{\Omega}$ . Then the problem*

$$\begin{aligned} \frac{\partial u}{\partial t} &= \sum_{i,j=1}^m \frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u}{\partial x_j} \right) + f(u) \text{ for } (x, t) \in \Omega \times J; \\ u(x, 0) &= \bar{u}(x) \quad x \in \bar{\Omega}; \\ \frac{\partial u}{\partial \mathbf{n}}(x, t) &= 0 \text{ for } (x, t) \in \partial\Omega \times J, \end{aligned} \tag{3.15}$$

has a unique solution  $(u_1(x, t), u_2(x, t), \dots, u_n(x, t)) \in \bar{\Omega} \times J$  where  $J = (0, T]$  for any  $T > 0$ . If an a priori estimate  $u(x, t) \leq K$  for  $0 \leq t \leq T, x \in \bar{\Omega}$  can be established, where  $K$  is independent of  $T$ , then the solution exists for all  $t \geq 0$ .

Therefore the long time behavior of the solution for this replicator diffusion equation is stated in the following theorem.

**Theorem 3.3.5.** *Let the ordinary differential equation (3.6) have an exponentially stable stationary state with all components positive (stable polymorphism). Assume the bounded nonlinear payoff function  $\pi(p)$  satisfies the following two conditions*

$$[(\nabla\pi(p)^T \cdot p)_i - (\nabla\pi(p)^T \cdot p)_j] [\pi_i(p) - \pi_j(p)] \geq 0, \forall i \neq j; \quad (3.16)$$

$$x^T \nabla\pi(p)x \leq 0, \forall x \in R^n \text{ satisfies } \sum_{i=1}^n x_i = 0, \quad (3.17)$$

where the equality in (3.17) holds iff  $x = 0$ . Then the solution of the replicator diffusion equation (3.7) with Neumann boundary condition (3.9) and initial condition (3.14) converges uniformly in  $\Omega$  towards a spatially homogeneous (i.e., constant) function, which is a stationary state of the ordinary differential equation (3.6).

*Proof.* Denote the state space of the variable to be

$$S = \{p \in R^n : \sum_{i=1}^n p_i = 1, p_i \geq 0, i = 1, \dots, n\}.$$

Since the population density function  $p$  is bounded above by 1, then the existence and uniqueness of the solution of the replicator diffusion equation is guaranteed by Lemma 3.3.4. Also by the boundedness of the solution  $p(x, t)$ , the path  $\{p(\cdot, t), t \geq t_0 > 0\}$  is compact in  $C^1(\Omega \rightarrow R^n)$ , and  $C^2(\Omega \rightarrow R^n)$  as well.

Define the average payoff function  $W(p) : S \rightarrow R$  where  $W(p) = p^T \pi(p)$  and it's a Lyapunov function for (3.6) because

$$\begin{aligned} \frac{dW(p(t))}{dt} &= p_t^T (\pi(p) + \nabla\pi(p)^T \cdot p) \\ &= (\pi(p)^T P - (p^T \pi(p))p^T)(\pi(p) + \nabla\pi(p)^T \cdot p) \\ &= [\pi(p)^T P\pi(p) - (p^T \pi(p))^2] + [p^T \nabla\pi(p)P\pi(p) - p^T \nabla\pi(p)p \cdot p^T \pi(p)], \end{aligned} \quad (3.18)$$

where

$$\pi(p)^T P\pi(p) - (p^T \pi(p))^2 = \sum_{i=1}^n p_i \pi_i^2(p) - \left( \sum_{i=1}^n p_i \pi_i(p) \right)^2 \geq 0$$

by the Cauchy-Schwartz inequality and

$$\begin{aligned}
& p^T \nabla \pi(p) P \pi(p) - p^T \nabla \pi(p) p \cdot p^T \pi(p) \\
&= \sum_{i=1}^n p_i (\nabla \pi(p)^T \cdot p)_i \pi_i(p) - \sum_{i=1}^n p_i (\nabla \pi(p)^T \cdot p)_i \sum_{j=1}^n p_j \pi_j(p) \\
&= \sum_{i=1}^n p_i (\nabla \pi(p)^T \cdot p)_i \pi_i(p) \sum_{j=1}^n p_j - \sum_{i=1}^n p_i (\nabla \pi(p)^T \cdot p)_i \sum_{j=1}^n p_j \pi_j(p) \\
&= \frac{1}{2} \left\{ \sum_{i \neq j} p_i p_j [(\nabla \pi(p)^T \cdot p)_i \pi_i(p) + (\nabla \pi(p)^T \cdot p)_j \pi_j(p) - (\nabla \pi(p)^T \cdot p)_i \pi_j(p) - \pi_i(p) (\nabla \pi(p)^T \cdot p)_j] \right\} \\
&= \frac{1}{2} \left\{ \sum_{i \neq j} p_i p_j [(\nabla \pi(p)^T \cdot p)_i - (\nabla \pi(p)^T \cdot p)_j] [\pi_i(p) - \pi_j(p)] \right\} \geq 0.
\end{aligned}$$

The above two nonnegative quantities vanish if  $\pi_i(p) = \pi_j(p), \forall i \neq j$ , that is,  $p$  is stationary since  $p_t = 0$  in (3.6).

Consider a particular solution  $p(x, t)$  of the initial value problem (3.8), (3.9) and (3.14). Denote a spatial Lyapunov function  $w : \Omega \times R^+ \rightarrow [-K, K]$  where  $K = \max\{|W(p(x, t))| : (x, t) \in \Omega \times R^+\}$  by

$$w(x, t) = W(p(x, t)). \quad (3.19)$$

Then take the partial derivatives, we obtain

$$\begin{aligned}
w_{x_i} &= p_{x_i}^T \pi(p) + p^T \nabla \pi(p) p_{x_i} \\
w_{x_i x_i} &= p_{x_i x_i}^T \pi(p) + 2p_{x_i}^T \nabla \pi(p) p_{x_i} + p^T \nabla \pi(p) p_{x_i x_i} \\
\Delta w &= \Delta p^T \pi(p) + p^T \nabla \pi(p) \Delta p + 2 \sum_{i=1}^N p_{x_i}^T \nabla \pi(p) p_{x_i} \\
w_t &= p_t^T \pi(p) + p^T \nabla \pi(p) p_t.
\end{aligned}$$

It follows that for  $(x, t) \in \Omega \times R^+$ ,

$$w_t - \Delta w = (p_t - \Delta p)^T (\pi(p) + \nabla \pi(p)^T p) - 2 \sum_{i=1}^N p_{x_i}^T \nabla \pi(p) p_{x_i} \quad (3.20)$$

where  $(p_t - \Delta p)^T (\pi(p) + \nabla \pi(p)^T p)$  is in the same form as the right hand side term of (3.18) and for  $(x, t) \in \partial \Omega \times R^+$ ,

$$\frac{\partial w}{\partial \mathbf{n}} = 0. \quad (3.21)$$

We define the spatial average of  $w(x, t)$  to be

$$I(t) = \int_{\Omega} w(x, t) dx / |\Omega|. \quad (3.22)$$

Then integrate (3.20) we obtain

$$\int_{\Omega} w_t dx - \int_{\Omega} \Delta w dx = \int_{\Omega} [(p_t - \Delta p)^T (\pi(p) + \nabla \pi(p)^T p) - 2 \sum_{i=1}^N p_{x_i}^T \nabla \pi(p) p_{x_i}] dx \geq 0.$$

Thus from the Neumann boundary condition we get

$$\frac{dI(t)}{dt} = \left( \int_{\Omega} w_t dx - \int_{\Omega} \frac{\partial w}{\partial \mathbf{n}} dS \right) / |\Omega| = \left( \int_{\Omega} w_t dx - \int_{\Omega} \Delta w dx \right) / |\Omega| \geq 0. \quad (3.23)$$

Since  $I(t)$  is bounded above by  $K = \max\{|W(p(x, t))| : (x, t) \in \Omega \times R^+\}$ , therefore it converges from below to some  $\bar{I}$ .

Based on Lemma 3.3.2, for every  $t, \tau \geq 0$  the function  $w(x, t + \tau)$  is bounded below by the solution  $v(x, t, \tau)$  of the equation

$$\begin{aligned} v_t - \Delta v &= 0; \\ v(x, t, 0) &= w(x, t); \\ \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial \Omega} &= 0. \end{aligned} \quad (3.24)$$

From Lemma 3.3.3, the solution  $v(x, t, \tau)$  approximates a constant as  $\tau \rightarrow \infty$ , that is,

$$v(x, t, \tau) \rightarrow I(t). \quad (3.25)$$

Thus

$$\liminf_{\tau \rightarrow \infty} w(x, t + \tau) \geq I(t).$$

That is,

$$\bar{w}(x) = \liminf_{t \rightarrow \infty} w(x, t) \geq \lim_{t \rightarrow \infty} I(t) = \bar{I}. \quad (3.26)$$

Next we show that

$$\lim_{t \rightarrow \infty} w(x, t) = \bar{I}. \quad (3.27)$$

Assume there is a point  $x_0 \in \Omega$  and a sequence  $\{t_k\}_{k=1}^{\infty}$  with  $t_k \rightarrow \infty$  such that  $w(x_0, t_k) \rightarrow I_1 > \bar{I}$ . Then since  $w(x, t)$  is bounded on  $\Omega \times R^+$ , we choose a subsequence  $\{t_{k_j}\}_{j=1}^{\infty}$  such that  $w(x, t_{k_j}) \rightarrow w_1(x)$  uniformly in  $\Omega$ . Thus  $w_1(x_0) = I_1 > \bar{I}$ . From (3.26) we know  $w_1(x) \geq \bar{I}$ , but  $\int_{\Omega} w_1(x) dx / |\Omega| = \lim_{k \rightarrow \infty} \int_{\Omega} w(x, t_{k_j}) dx / |\Omega| = \lim_{k \rightarrow \infty} I(t_{k_j}) = \bar{I}$ . Contradiction!

Thus  $w(\cdot, t)$  converges to a constant  $\bar{I}$  for all  $x \in \Omega$  as  $t \rightarrow \infty$ . Then it follows that

$$w_t \rightarrow 0 \quad (3.28)$$

and

$$\Delta w \rightarrow 0. \quad (3.29)$$

From (3.19) we therefore have

$$[(p_t - \Delta p)^T(\pi(p) + \nabla\pi(p)^T p) - 2 \sum_{i=1}^N p_{x_i}^T \nabla\pi(p) p_{x_i}] \rightarrow 0. \quad (3.30)$$

Since  $(p_t - \Delta p)^T(\pi(p) + \nabla\pi(p)^T p) \geq 0$  and  $\sum_{i=1}^N p_{x_i}^T \nabla\pi(p) p_{x_i} \leq 0$ , then

$$(p_t - \Delta p)^T(\pi(p) + \nabla\pi(p)^T p) \rightarrow 0 \quad (3.31)$$

and

$$\sum_{i=1}^N p_{x_i}^T \nabla\pi(p) p_{x_i} \rightarrow 0. \quad (3.32)$$

For any sequence  $\{t_k\}_{k=1}^\infty$  with  $t_k \rightarrow \infty$  there is a subsequence  $\{t_{k_j}\}_{j=1}^\infty$  such that  $p(x, t_{k_j})$  converges in  $C^2$ . From the fact that  $\sum_{i=1}^N p_{x_i}^T \nabla\pi(p) p_{x_i} \leq 0$  and (3.32), we know that all first derivatives of  $p(\cdot, t_{k_j})$  approach 0, thus  $p(\cdot, t_{k_j})$  goes to a constant. By (3.18) and (3.32) we know this constant is a stationary solution of (3.6).  $\square$

**Remark 3.3.6.** For linear game with  $n$  strategies, that is,  $\pi(p) = Ap$ , the left hand side of condition (3.16) is perfect square and condition (3.17) can be derived from the existence of the exponentially stable stationary state. However for nonlinear games,  $\nabla\pi(p)$  depends on local  $p$  and is not generally the constant matrix  $A$  any more. Then the existence of exponentially stable stationary state can no longer derive condition (3.17) as that in linear games. So theorem 3.3.1 is a special case of theorem 3.3.5, and theorem 3.3.5 still holds without the condition for existence of exponentially stable stationary state.

### 3.4 Examples and Numerical Simulation Results

For the system described in section 3.3, there is no closed form solution for it, thus we investigate the finite difference algorithms for studying the dynamics of spatially distributed population based on the replicator diffusion equations given appropriate initial and boundary conditions. The convergence and stability analysis of the finite difference schemes will also be studied as followed.

We consider the parabolic system in a two-dimensional space. To construct the finite-difference methods in a bounded rectangle region, we take a uniform subdivision of the region  $\Omega = [0, A] \times [0, B]$

with grid points  $(x_i, y_j) = (ih, jh), i = 0, \dots, I, j = 0, \dots, J$  where  $I = A/h, J = B/h$ . We also use a uniform subdivision of the time interval  $[0, T]$  with time levels  $t_k = k\Delta t, k = 1, \dots, K$  where  $K = T/\Delta t$ . Then the approximate of the solution  $p(x, t)$  in a three dimension region  $\Omega \times [0, T]$  is denoted by  $p_{i,j}^k = (p_{1,i,j}^k, \dots, p_{n,i,j}^k)$ . We can number the two-dimensional grid in a natural way, that is, number consecutively from left to right starting from the bottom row, from  $m = 0, 1, \dots, M$  where  $M = (I + 1)(J + 1) - 1$ . Thus we have the following

$$p_{s,m}^k = p_{s,i,j}^k, \quad s = 1, \dots, n; m = i + j(I + 1); i = 0, \dots, I; j = 0, \dots, J. \quad (3.33)$$

Denote  $p_{i,j}^k = (p_{1,i,j}^k, \dots, p_{n,i,j}^k)$ . We use the following notation to simplify the forward differences in time, the five point central difference approximation of the Laplacian in two dimensions:

$$\begin{aligned} \partial_k \phi &= (\phi^k - \phi^{k-1})/\Delta t, \\ \Delta_h \psi &= (\psi_{i,j-1} + \psi_{i,j+1} + \psi_{i-1,j} + \psi_{i+1,j} - 4\psi_{i,j})/h^2. \end{aligned} \quad (3.34)$$

Next we discuss three two-dimensional linear schemes for the replicator diffusion equation in (3.8), (3.9) and (3.14). First of all, for  $k = 1, \dots, K$  and  $i = 0, \dots, I, j = 0, \dots, J$  find  $p_{s,i,j}^k, s = 1, \dots, n$  such that

$$\partial_k p_{s,i,j}^k = \Delta_h p_{s,i,j}^{k-1} + p_{s,i,j}^{k-1} \left( \pi_s(p_{i,j}^{k-1}) - p_{i,j}^{k-1} \pi(p_{i,j}^{k-1}) \right) \quad (3.35)$$

with initial conditions

$$p_{i,j}^0 = p_{0,i,j} = p_0(x_i, y_j). \quad (3.36)$$

From a computational point of view, the most convenient way is to use the explicit method in (3.35) to approximate the solution of the continuous system (3.8). However, this method imposes a severe restriction on the step size for numerical stability. An alternative approach is the semi-implicit method in the sense that the Laplacian operator is replaced by an implicit approximation while the reaction term is approximated by an explicit expression. That is,

$$\partial_k p_{s,i,j}^k = \Delta_h p_{s,i,j}^k + p_{s,i,j}^{k-1} \left( \pi_s(p_{i,j}^{k-1}) - p_{i,j}^{k-1} \pi(p_{i,j}^{k-1}) \right). \quad (3.37)$$

The third way is the implicit method that requires some iterative scheme for the computation of numerical solutions and leads to a system of nonlinear algebraic equations which preserves the analogous qualitative property of the continuous system. The iterative process comes from the following

$$\partial_k p_{s,i,j}^k = \Delta_h p_{s,i,j}^k + p_{s,i,j}^k \left( \pi_s(p_{i,j}^k) - p_{i,j}^k \pi(p_{i,j}^k) \right). \quad (3.38)$$

This method can be seen in (25). But its main requirements are the mixed quasi-monotone property in the replicator reaction term and the existence of a pair of coupled upper and lower solutions which are hard to be satisfied.

We consider two examples of nonlinear games with two strategies  $C$  and  $D$ . The game dynamic model of the nonlinear games is

$$\begin{aligned}\frac{df}{dt} &= f(P_C(f, h) - fP_C(f, h) - hP_D(f, h)) \\ \frac{dh}{dt} &= h(P_D(f, h) - fP_C(f, h) - hP_D(f, h)),\end{aligned}\tag{3.39}$$

where  $f$  is the frequency of strategy  $C$  players,  $h$  is the frequency of strategy  $D$  players, and  $f + h = 1$  for each  $t$ . Assume that the fractional fixed point  $f^*$  for the replicator equation in (3.39) is asymptotically stable, and evolutionarily stable strategy as well. We further suppose that the payoff functions satisfy the conditions of the theorem 3.3.5 and from this theorem, we know that the solution of the adapted replicator diffusion equation

$$\begin{aligned}\frac{\partial f}{\partial t} &= \Delta f + f(P_C(f, h) - fP_C(f, h) - hP_D(f, h)) \\ \frac{\partial h}{\partial t} &= \Delta h + h(P_D(f, h) - fP_C(f, h) - hP_D(f, h)) \\ \frac{\partial f}{\partial \mathbf{n}} \Big|_{\partial \Omega} &= 0 \\ \frac{\partial h}{\partial \mathbf{n}} \Big|_{\partial \Omega} &= 0\end{aligned}\tag{3.40}$$

given any piecewise smooth initial condition in the bounded compact region  $\Omega$  goes to a spatially homogeneous stationary point of (3.39) uniformly as  $t \rightarrow \infty$ . We can also numerically verify this conclusion on the long time behavior of the adapted replicator diffusion equations. From the constraint  $f + h = 1$ , the replicator diffusion equations can be reduced to

$$\begin{aligned}\frac{\partial f}{\partial t} &= \Delta f + f(P_C(f, 1 - f) - fP_C(f, 1 - f) - (1 - f)P_D(f, 1 - f)) \\ \frac{\partial f}{\partial \mathbf{n}} \Big|_{\partial \Omega} &= 0.\end{aligned}\tag{3.41}$$

Using the explicit finite difference method, we obtain the following iterative process

$$\begin{aligned}\partial_k f_{s,i,j}^k &= \Delta_h f_{s,i,j}^{k-1} + f_{s,i,j}^{k-1} \left( P_C(f_{s,i,j}^{k-1}, 1 - f_{s,i,j}^{k-1}) - f_{s,i,j}^{k-1} P_C(f_{s,i,j}^{k-1}, 1 - f_{s,i,j}^{k-1}) \right. \\ &\quad \left. - (1 - f_{s,i,j}^{k-1}) P_D(f_{s,i,j}^{k-1}, 1 - f_{s,i,j}^{k-1}) \right).\end{aligned}\tag{3.42}$$

After adding the boundary and initial conditions, the above formula can be simplified and it boils down to

$$\begin{aligned}\partial_k f_{s,i,j}^k &= \Delta_h f_{s,i,j}^{k-1} + f_{s,i,j}^{k-1} (1 - f_{s,i,j}^{k-1}) \left( P_C(f_{s,i,j}^{k-1}, 1 - f_{s,i,j}^{k-1}) - P_D(f_{s,i,j}^{k-1}, 1 - f_{s,i,j}^{k-1}) \right) \\ B^{(k)}[f_{s,i,j}^k] &= 0 \\ f_{s,i,j}^0 &= p_{0,i,j},\end{aligned}\tag{3.43}$$

where  $B^{(k)}[f_{s,i,j}^k]$  is a suitable approximations of the Neumann boundary operator  $B = \frac{\partial}{\partial n}$  at the  $k$ -th time step. Alternatively we apply the semi-implicit approach and similarly we have the iterative procedure as follows

$$\begin{aligned}\partial_k f_{s,i,j}^k &= \Delta_h f_{s,i,j}^k + f_{s,i,j}^{k-1} (1 - f_{s,i,j}^{k-1}) \left( P_C(f_{s,i,j}^{k-1}, 1 - f_{s,i,j}^{k-1}) - P_D(f_{s,i,j}^{k-1}, 1 - f_{s,i,j}^{k-1}) \right) \\ B^{(k)}[f_{s,i,j}^k] &= 0 \\ f_{s,i,j}^0 &= p_{0,i,j}.\end{aligned}\tag{3.44}$$

The semi-implicit scheme can be expressed as  $IJ$  linear equations at each time step  $t_k, k = 1, \dots, K$ . It's computationally efficient to apply the conjugate gradient method to solve the linear system instead of the solving scheme using high dimensional matrix decomposition. The semi-implicit scheme is a stable finite-difference approximation of the solution of the parabolic differential equation.

In the numerical simulations, the initial data is set to be  $p_{0,i,j} = 0.02$  if  $|x_i| \leq \frac{A}{2}, |y_j| \leq \frac{A}{2}$ ,  $p_{0,i,j} = 1$  if  $(x_i - \frac{3A}{4})^2 + (y_j - \frac{3A}{4})^2 \leq (\frac{A}{4})^2$  and 0.3 otherwise where  $A$  means the side length of the squared region. All codes were implemented in MATLAB 7.14 (R2012a).

### 3.4.1 Constructed Example

Consider a nonlinear game that is artificially constructed with the following nonlinear payoff functions for two strategies  $C$  and  $D$ :

$$\begin{aligned}P_C &= -\log \left( f + \frac{1}{2} \right) \\ P_D &= \log \left( f + \frac{1}{2} \right).\end{aligned}\tag{3.45}$$

From the categorization criteria of nonlinear games in (40), we know that this game is a Snowdrift nonlinear game. And the unique Nash equilibrium is  $f^* = \frac{1}{2}$ , which is both evolutionarily stable and asymptotical stable. It can be verified that the payoff functions in this game satisfy the conditions (3.16)

and (3.17) in theorem 3.3.5. Then theoretically the solution of the corresponding adapted converges uniformly to a stationary state of (3.39). We also verify this conclusion from the simulation study and the results are provided in Figure 3.1 and 3.2.

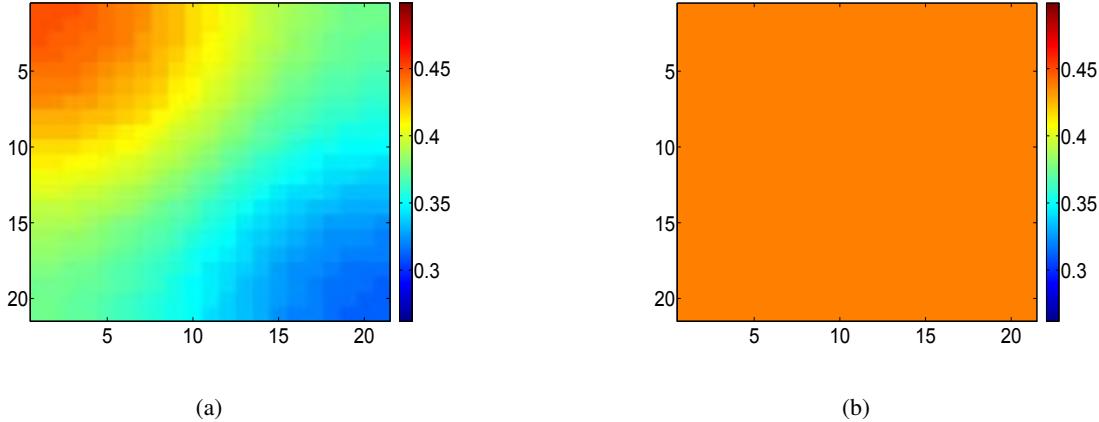


Figure 3.1 Two dimensional approximates for strategy  $C$  frequency at different time points. (a)  $t = 0.2$ . (b)  $t = 1.5$ .

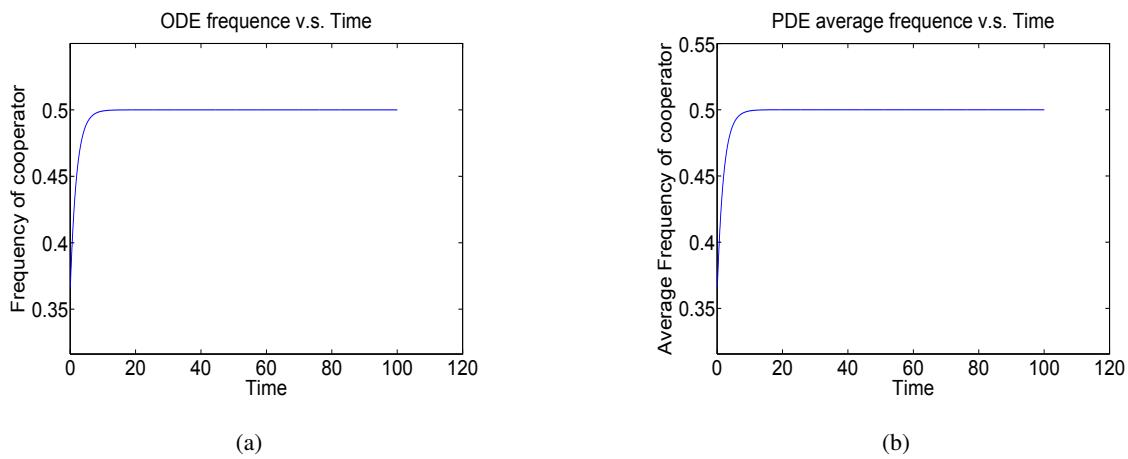


Figure 3.2 Long time behavior of the average frequency of strategy  $C$  in adapted replicator diffusion equation and comparison of it and replicator dynamics. (a) Replicator dynamics. (b) Adapted replicator diffusion equation.

Figure 3.1 suggests that the frequencies of strategy  $C$  at all locations are becoming homogeneous as the time increases. That is, the cooperator frequencies tend to the same constant uniformly during the initial time period. This indicates that the diffusion rate 1 is large and the diffusion effects is significant

which dominates the mechanism of the replicator diffusion equation by comparing to the replicator source term. As the frequencies of all individuals at the grids converge to the constant homogeneously, then there is nearly no gap between the frequency of each grid and that of its neighbors, and thus the diffusion effects become ignorable and the reaction source effects take over the dominance and determine the mechanism of the equation. From then on, the diffusion term will be approximately zero and the adapted replicator diffusion equations behave the same as the replicator dynamics, that is, the overall solution of the replicator diffusion equation for every spatial location is almost identical to the solution in the replicator dynamics. This is consistent with the patterns shown in Figure 3.2.

In Figure 3.2, the first plot represents the solution curves of the replicator dynamics with respect to time and the second plot gives us the average strategy  $C$  frequencies of the adapted replicator diffusion equations. The average frequency of the parabolic systems converge to a constant 0.5 which is the equilibrium of the games. That means the mechanism of the replicator dynamics and that of adapted replicator diffusion equations are similar in a well-mixed population. We can therefore conclude that the frequencies of the strategy  $C$  players in the parabolic equations will converge uniformly to the spatial homogeneous solution of the nonlinear replicator dynamics which is an evolutionarily stable state.

### 3.4.2 Yeast Example

Consider another example of nonlinear game in a yeast system proposed in (26; 31). There are two types of strains—cooperator and cheater strains (defector) and Gore et al. found that the interaction between these two strains could be either snowdrift game or prisoner’s dilemma, and the nature of this cooperation game depends on the parameters interpreted as the glucose concentration and the histidine concentration levels. That is, the cooperator strain in the experiments is histidine auxotroph and limiting the histidine concentration will experimentally increase the cost of cooperation. If the rare strategies which will often interact with the opposite strategy do comparatively well under certain levels of glucose and histidine, then there will be coexistence between cooperation and defection and the game becomes snowdrift game. However if the histidine concentration is low and the glucose concentration is high, the cheating strains will always do better and the defectors outgrow the cooperators which leads to the extinction of cooperators and the game will be prisoner’s dilemma game. These can be seen in Figure 3.3.

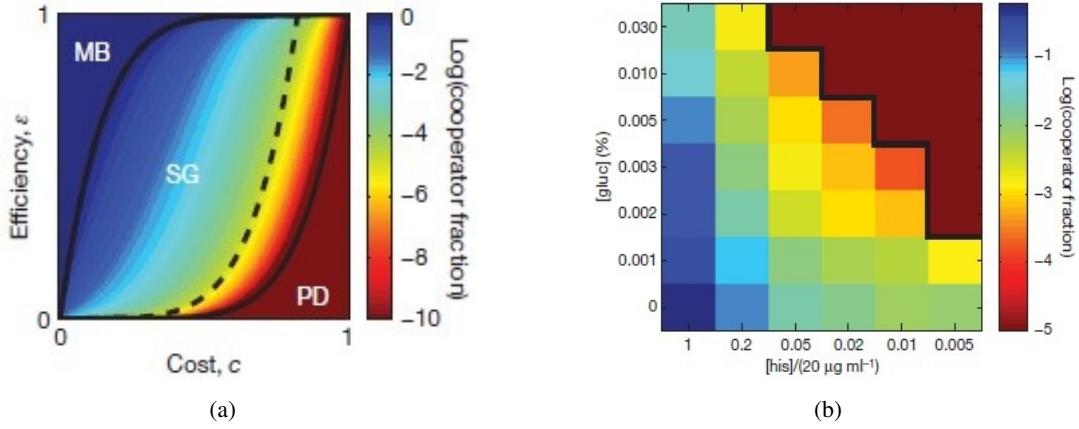


Figure 3.3 Game theory models of cooperation in yeast system. (a) A model of cooperation with experimentally measured concave benefits yields a central region of parameter space that is a snowdrift game. (b) Fraction of cooperators at equilibrium as a function of the glucose and histidine concentrations.

They also measured the growth rate as a function of glucose concentration and cooperator strain fractions and then derived the nonlinear payoff functions for these two types of strains. Thus they imposed the capture efficiency  $\epsilon$ , cost  $c$  and the power of  $\alpha$  onto the benefit terms and fitted the payoff functions with respect to those parameters and cooperator frequency. Thus the payoff functions of cheater or defector and cooperator are the following

$$P_C = [g + (1 - g)(\epsilon + f(1 - \epsilon))]^\alpha - c \quad (3.46)$$

$$P_D = [g + (1 - g)f(1 - \epsilon)]^\alpha,$$

where  $f$  is the frequency of cooperator strains in the yeast population. Huang and Wu established the spatial homogeneous game dynamic model for the yeast system (31). They also established the fact that the fractional fixed point  $f^*$  for the replicator equation in (3.39) of the yeast game is asymptotically stable, and evolutionarily stable strategy as well. However the conditions of the theorem 3.3.5 are not completely satisfied for the nonlinear games with two different set of parameters which suggest it's not sufficient to prove that the solution of the adapted replicator diffusion equation given any piecewise smooth initial condition in the bounded compact region  $\Omega$  goes to a stationary state uniformly as  $t \rightarrow \infty$ . Even though the convergence of the solution cannot be guaranteed theoretically, we numerically justify the asymptotical behavior of the solution to the adapted replicator diffusion equation in the yeast system.

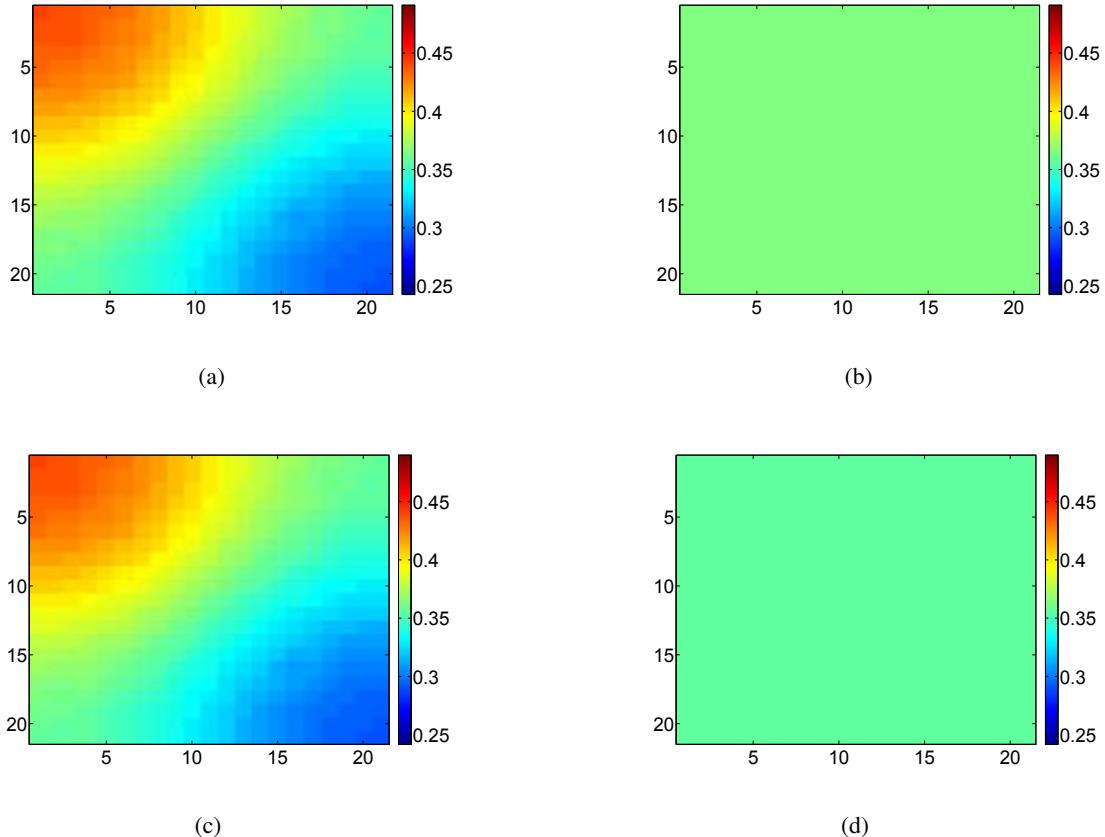


Figure 3.4 Two dimensional approximates for the cooperator strain frequency using different parameter sets at different time points. (a)  $A = 1, a = 1, g = 0.00, c = 0.0023, T = 5000, I = 20$  at  $t = 0.2$ . (b)  $A = 1, a = 1, g = 0.00, c = 0.0023, T = 5000, I = 20$  at  $t = 1.5$ . (c)  $A = 1, a = 1, g = 0.03, c = 0.0254, T = 5000, I = 20$  at  $t = 0.2$ . (d)  $A = 1, a = 1, g = 0.03, c = 0.0254, T = 5000, I = 20$  at  $t = 1.5$ .

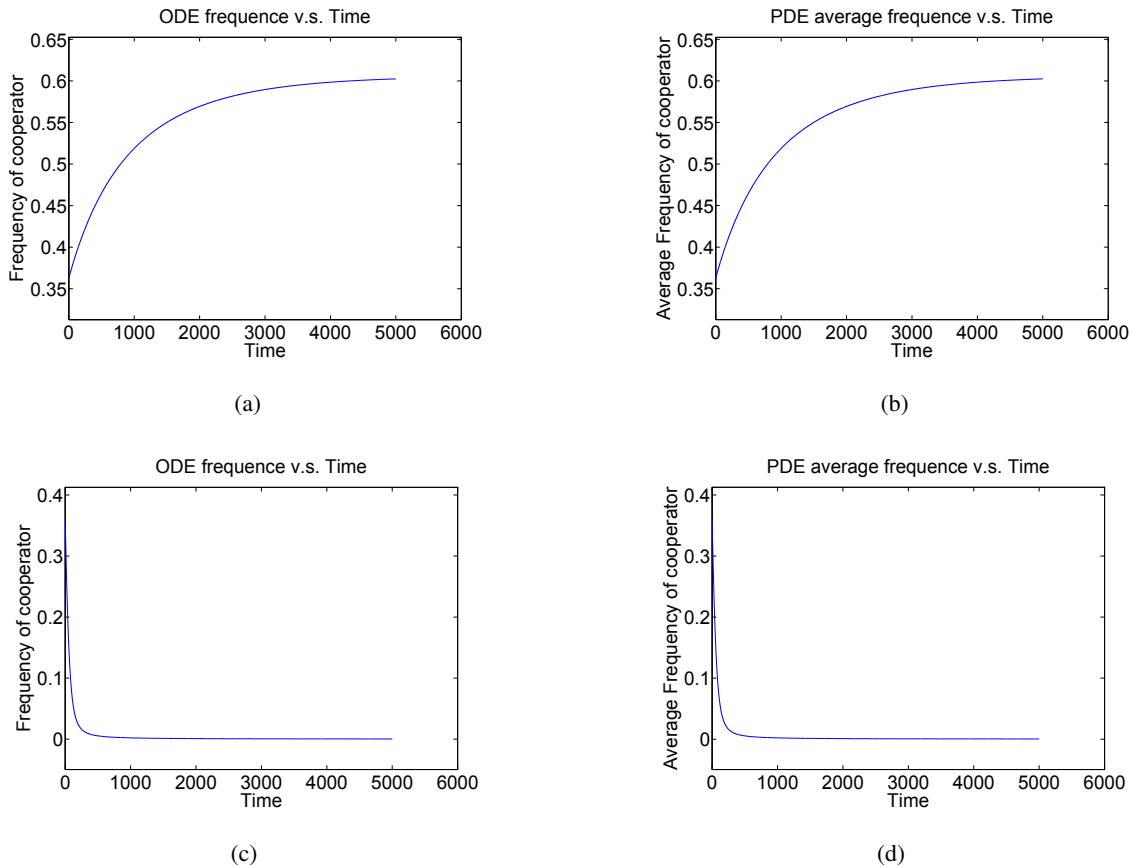


Figure 3.5 Long time behavior of the average frequency of cooperators in adapted replicator diffusion equations and comparison of replicator dynamics and adapted replicator diffusion equations . (a)-(b) Snowdrift game where the parameter set is  $A = 1, a = 1, g = 0.00, c = 0.0023, T = 5000, I = 20$ , and the experimental equilibrium is 0.6. (c)-(d) Prisoner's Dilemma where the parameter set is  $A = 1, a = 1, g = 0.03, c = 0.0254, T = 5000, I = 20$ , and the experimental equilibrium is 0.

In the numerical schemes, the frequency density of cooperators in (3.44) is calculated with respect to various parameter sets that were chosen in the yeast experiment of Gore et al. Associated with the computer implementation, the results of the computer experiments are presented in Figure 3.4 and 3.5.

From the plots in Figure 3.4, we can see that there are two different parameter sets and they correspond to two different types of games: Snowdrift game and Prisoner Dilemma. Even though the nature of the game are different, they all present the same pattern that the frequencies at all locations are becoming homogeneous. And similar to the constructed example, the cooperator frequencies tend to a constant uniformly which is the initial average frequency. It tells us that the diffusion rate is large and the diffusion effect is dominating the mechanism of the replicator diffusion equation initially. This may explain the well-mixedness of some biological systems caused by the diffusion or dispersal effects from internal or external forces. When the frequencies of individuals at all locations converge to the same constant, the diffusion term approximates zero and the reaction source effect starts to dominate and decides the mechanism of the dynamics. Afterwards, the spatial term is almost zero and the adapted replicator diffusion equations behave like the replicator dynamics, that is, the behavior of the solution at every spatial location in the replicator diffusion equation is almost the same that in the replicator dynamics. This agrees with the curves shown in Figure 3.5.

The first column in Figure 3.5 represents the solution curves of the replicator dynamics with respect to time under two different sets of parameters and the second column gives us the average cooperator frequencies of the adapted replicator diffusion equations in corresponding two types of games. The spatial average of the solution converges to some constant for each parameter set where the constant is approximately the equilibria of corresponding game. Also we compare the cooperator frequencies in replicator dynamics and the average cooperator frequencies in corresponding adapted replicator diffusion equations in each row of Figure 3.5, we can see that there is no much difference between the two curves. That means the mechanism of the replicator dynamics is similar to that of adapted replicator diffusion equations. We further conclude that the frequencies of the cooperator strain in the parabolic equations will converge uniformly to the spatial homogeneous stationary state of the replicator dynamics in yeast game. So in the perspective of computation, we get some important clues to the behavior of two yeast strains in the spatially extended environment and this also gives a reasonable geological interpretation for the uniform spatial convergence in theorem 3.3.5.

### 3.5 Discussion and Conclusions

In this paper, we have discussed the equivalent conditions of the evolutionarily stable strategy for a general class of replicator dynamics. We also establish the theory for the long time behavior of the solution in the adapted replicator diffusion equations which incorporates the spatial effects. Finally the construct artificial system and the yeast cooperation system in Gore et al.(2009) are tested out to verify the long time behavior of the adapted replicator diffusion equations generated from the game dynamics.

Although it's difficult to find the necessary and sufficient conditions of the existence of evolutionarily stable strategies for the direct form of the payoff functions of all species in a population game based on the replicator dynamics, we derive the equivalent conditions of ESS from the differential behavior of the payoff functions at the stationary state indirectly using the Taylor expansion idea. The coefficients of two Taylor series are compared based on the definition of the ESS and this yields to infinite series of inequalities involving the infinite differential properties of the payoff functions at the stationary state. We find that the stronger version of first two inequalities are necessary to the existence of a regular ESS which is an refinement of ESS. Also these two inequalities are necessary and sufficient for an ESS when the game is a linear game.

As for the general replicator dynamics of multiple species in a population, the spatial effects determined by the environment are introduced and the adapted replicator diffusion equation are then generated from the general replicator equations by adding the same diffusion term for all species. Based on the lemmas about the convergence of simple heat equation and the existence and uniqueness of a general class of parabolic equations, we prove the uniform convergence of the solution of the adapted replicator diffusion equation to the exponentially stable stationary state. This suggests that the cooperation and competition of several species in a spatially distributed population finally reaches an equilibrium which is the same as those well mixed population in the same payoff functions situation from the biological point of view.

The constructed game and the yeast game proposed by Gore et al. are classical population games within two populations. In the yeast game, Gore et al. examined the wild type strain (cooperator) and the mutant strain (defector). The authors derived the payoff functions for this yeast system which was a nonlinear game. The dynamics and equilibrium stability of this game were investigated by Huang

and Wu (2012) and the results fitted the experimental ones very well. The replicator equations for the artificial game, and especially for the yeast game, are adapted by adding diffusion terms because of the nature of strains digesting the dispersal glucose in the environment. And we use the semi-implicit algorithm to generate the numerical solutions of the two adapted equations. The contour plots of the frequencies of the strategy  $C$  are plotted, and they indicates that it is the diffusion effect that drives the behaviors of the solutions initially. The latter process are then governed by the replicator source term and this is justified by the contour plots of cooperator frequencies and the comparison plots of the frequencies under the well-mixed replicator dynamics and the average frequencies under the adapted replicator diffusion equations for different types of nonlinear games. The payoff functions in yeast game don't satisfy both conditions in theorem 3.3.5, but the numerical experiments still indicates the long time behavior of the solution, which suggests that the conditions in theorem 3.3.5 may be stronger for the asymptotic convergence of the solution.

The present modeling study shows the asymptotical behavior of the adapted replicator diffusion equations under the same diffusion rate situation. It also filling one with the interest in evolutionary game theory and its applicators from both mathematical and biological perspective. Our future work will consider the behavior of the solution to the parabolic systems with weaker conditions than those in theorem 3.3.5 and more general diffusion effects which means the dispersal rate of the species in a population vary or be more general operator rather than Laplacian, and the methodology of parameter estimation in such replicator diffusion equations.

### Acknowledgments

The authors would like to thank Prof. Yuan Lou for many insightful suggestions in the final stage of this work. The authors would also like to thank Prof. Wen Zhou for providing the Matlab sample code of the general semi-implicit algorithms for solving the parabolic equations and giving valuable suggestions.

## CHAPTER 4. PARAMETER ESTIMATION IN DYNAMICAL METAGENOMIC MODEL

### Abstract

Many biological systems and evolutionary processes can be described by a set of temporal mathematical models, especially evolutionary game models. There is sparse literature for the estimation of parameters in the game theoretical models, and it's challenging to infer the parameters from raw data. We propose a specific model for the temporal evolution of mixed populations, such as those queried in modern metagenomics. The problem for identifiability of the parameters in the model is addressed. A simulation study is implemented to test the feasibility and acceptability of the procedure. We will apply the proposed methodology and techniques to a microbial data set from gut communities of mice. It is expected that the proposed modeling and estimation approaches be used for a variety of evolutionary game studies.

### 4.1 Introduction

In biological systems, it is often of great interest to predict the future prosperity, stability, dominance or extinction of an individual organism or species, and to describe and interpret the changes among them in a population as well. With the advent of modern sequencing technology, it is possible to query microbial community dynamics at previously unattainable scales, many species at frequent sampling times. In the population-based modeling approaches, the internal states and phenotypic functions are assumed to be homogeneous across cells within each species. One of the main issues therein is how to account for interspecies interactions in a direct or indirect way for the prediction of community structure and functions. A review of static methods for inferring interspecies interactions and approaches for predicting community structure in a given condition and its dynamic change in response to perturbations

is presented (63). Our focus is on inference from temporal samples. For the sake of simplicity, we confine our discussion to well-mixed environments.

Advanced metagenomics techniques allow for the identification of member species in the community and estimates of their abundances or relative abundances. Sequence data can be segregated into bins representing distinct Operational Taxonomic Units (OTU), which may be individual organism or species, or groups of species that share similar genomes. Microbial relationships can be inferred from species abundance or relative abundance data. The network of microbial associations can be predicted in a way using the techniques that include similarity-based methods, regression-based and rule-based methods. (For a review, see (47)). Another type of model—stoichiometric metabolic networks can be used to model individual species/taxa in microbial communities of equilibrium. In pioneering work by Stoylar et al., flux balance analysis (FBA) was applied to analyze the syntrophic relationship between sulfate-reducing bacteria and methanogens (62). Klitgord and Segre extended the method of Stoylar et al. to explore interspecies metabolite exchanges for various pairwise combinations of seven bacterial species (51). To relate microbe abundance with external covariates, nonlinear regression models including bioclimatic models and their extended forms are used to predict species abundance as a function of environmental conditions and its variation across a landscape (46; 52). While most researchers are familiar for kinetic approaches which deal with absolute concentrations, thermodynamic models potentially work with relative concentrations and may provide more mechanistic predictions on the change of community structure across conditions. For example, Larowe et al. used thermodynamic modeling to evaluate and compare the driving forces responsible for microbial anaerobic methane oxidation across representative marine ecological sites and different consortia (53). Finally species may not be the most natural units to describe the diversity and dynamics of microbial communities, Taxonomic or phylogenetic distinctions among species can become obscure and the dynamics of microbial communities better explained by focusing on microbial traits. Thus trait-based models, which use differential equations to represent the rates of change of nutrients, organic pools, etc., can be used for the steady state and dynamic analyses of diverse ecosystems. For instance, Boskill et al. introduced a microbial community trait-based modeling framework to study the nitrification in the microbe-mediated nitrogen cycle (43).

Suppose species abundances or relative frequencies are sampled over time. The Lotka-Volterra type equations or replicator equations mechanistically model the intrinsic growth rates and interactions

among the species. The Lotka-Volterra (LV) models are appropriate kinetic models for species abundance data, but abundances in sequence data are often dominated by technical noise. As an alternative, the replicator equations can model microbial population dynamics for relative abundances. The replicator equation arises out of evolutionary game theory, which was designed to study evolutionary biology and population interactions. It considers a population of players interacting in a game and each player or individual has fixed strategies. All the players interact with all others and receive fitness or payoff, and success in the game is translated into reproductive success. In 1978 Taylor and Jonker introduced a differential equation for evolutionary game dynamics (64), and later Zeeman, Schuster, Hofbauer and Sigmund did a general analysis on this type of differential equations (48; 49; 65). Suppose there exist interactions among  $n$  strategies and the payoff for a player with strategy  $i$  when interacting with another player choosing strategy  $j$  is given by the parameter  $a_{ij}$ . Then the fully structured  $n \times n$  matrix  $A = [a_{ij}]$  is called the payoff matrix which characterizes the interaction of  $n$  strategies. Let  $f_i$  denote the relative frequency of strategy  $i$  players in a population. The expected payoff of a strategy  $i$  player can be formulated as  $P_i = \sum_{j=1}^n a_{ij} f_j$  and the average payoff for the whole population is given by  $\phi = \sum_{i=1}^n f_i P_i$ . Thus the replicator equations to model the population interactions are

$$\dot{f}_i = f_i(P_i - \phi) \quad i = 1, \dots, n. \quad (4.1)$$

The above equations are defined on a simplex  $S_n = \{\mathbf{f} \in R^n : \sum_{i=1}^n f_i = 1; f_i \geq 0, \forall i\}$ , and they describe pure selection among the  $n$  strategies without mutation.

In this chapter, we explore the use of the replicator equations for studying the interaction of species in a microbial community sample over time using next generation sequencing applied to metagenomics. Specifically, we propose a parsimonious mixture model to reduce the number of free parameters. We combine information across multiple biological replicates, using the Negative Binomial to account for excess biological variation. We use smoothing spline approximation to the replicator equation solutions to avoid repeated integration. Finally, we demonstrate the feasibility of our model by testing it on simulation data. We end by discussing future directions that will further improve our estimation procedure.

## 4.2 Motivation

Marino et al. (2014) sequenced the 16S ribosomal gene from DNA isolated from the fecal material of germfree mice colonized with the cecal contents of conventionally raised animals and modeled these time-series data to characterize the interactions among the members of complicated murine microbial communities (54). To characterize the complex microbial communities, they classified the 16S sequence data into operational taxonomic units (OTU) representing species or higher taxonomic levels. They then formulated and fitted a system of ordinary differential equations to the OTU relative-abundance temporal data. The relative magnitudes of the intrinsic growth rates and the interactions between 17 OTUs were quantified by the following generalized Lotka-Volterra equations

$$\frac{dX_i(t)}{dt} = \alpha_i X_i(t) \left(1 - \frac{X_i(t)}{K}\right) + X_i(t) \left( \sum_{\substack{j=1 \\ j \neq i}}^n \beta_{ij} X_j(t) \right), \quad i = 1, 2, \dots, n, \quad (4.2)$$

where  $\alpha_i \geq 0$ ,  $\sum_i X_i \leq 1$  and  $n = 17$ . They concluded that all the OTUs had growth rates significantly above zero and of the 136 pairs of interactions between the 17 OTUs, 91 were competitive, 22 were parasitic and 17 were ammensalistic, 4 were commensal, 2 were neutral and none were mutualistic which indicated the gut environment was highly competitive, and numerous populations kept the others from dominating the community.

Though the dynamic model (4.2) can quantify the interactions among OTUs, it is not appropriate to model the relative abundance dynamics. Instead the replicator equations in evolutionary game theory are a good candidate to demonstrate the temporal process of OTU relative abundances. Even though the replicator equations and Lotka-Volterra equations are equivalent (Hofbauer, 1998), replicator dynamics are best used to describe the evolution of species proportions in a population driven by a frequency dependent fitness comparison while the Lotka-Volterra equations are suitable for describing density or absolute abundance dynamics of interacting species populations. Furthermore, Marino et al. were imposing positivity constraints on the intrinsic growth rates  $\alpha_i$ , which may not be appropriate. Certainly, interaction parameters for OTUs will be misestimated if the sign of the growth rate is misspecified, resulting in incorrect inferences about the OTU relationships. For  $n$  OTUs in a microbial community,

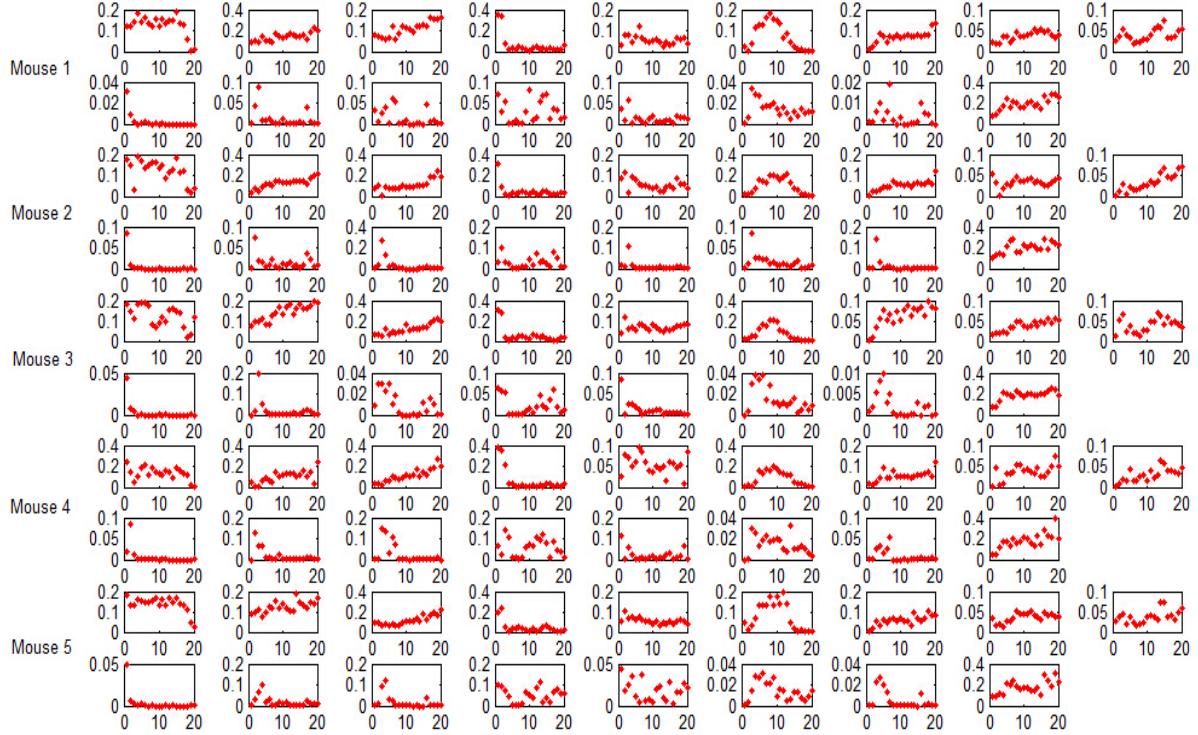


Figure 4.1 Observed temporal patterns for each OTU of 17 in each of 5 mice.

the direct application of the replicator dynamics with  $n$  equations which are the following

$$\dot{X}_i = X_i \left[ \sum_{j=1}^n a_{ij} X_j - \sum_{j=1}^n \sum_{k=1}^n a_{jk} X_j X_k \right] \quad i = 1, \dots, n, \quad (4.3)$$

is not suitable because of the identifiability issue, that is, the payoff matrix  $[a_{ij}]$  are not identifiable. Even though we can reparameterize the payoff matrix in terms of  $n(n - 1) = 272$  parameters which is large compared to the number of observations for the Marino data, where each OTU is observed at about 21 time points in 5 mice (Figure 4.1). Without any prior information on the relationship between the OTUs, we cannot limit the number of parameters to structure the dynamics. However, we observe similar temporal dynamics of certain OTUs across all mice in Figure 4.1. For instance, OTU 2, 3, 7, 8 and 17 (rare OTU) have the same increasing pattern, while OTU 4 and 10 display the same sudden disappearance soon after introduction. It is possible that these OTUs belong to related species, or families that share biological properties. In any case, it would be valuable to cluster OTUs into groups with similar dynamic patterns and to characterize the interactions among the clustered OTUs.

### 4.3 Model

It is assumed that there is a population of  $n$  OTUs ( $n$  usually is large, i.e.,  $n \geq 15$ ) in  $M$  mice, and the abundance of the OTUs in each mouse are measured at  $N$  time points  $t_j, j \in \{1, 2, \dots, N\}$ . Let  $y_{ijm}$  be the noisy count measurement of OTU  $i$  at time point  $t_j$  observed in replicate  $m$ . Then  $S_{jm} = \sum_{i=1}^n y_{ijm}$  is the observed sample size taken at  $t_j$  in mouse  $m$ . Two discrete probability distributions, the Poisson distribution or Negative-Binomial (NB) distribution are usually proposed to model sequence count data. However, the Poisson distribution is shown to be appropriate for sequence data only when no biological replicates are included (44; 55). When there are biological replicates, the sequence count data may exhibit more variability than expected with a Poisson distribution, i.e. the overdispersion phenomenon in (42). The NB distribution model for analysis of sequence data allows overdispersion and has been applied to RNA-seq data analysis. We consider the NB distribution in this article.

#### 4.3.1 Identifiability Analysis

To analyze the identifiability of the parameters of the payoff matrix  $A = [a_{ij}]$  in equation (4.1), the direct test method in (56) is implemented. That is, for the general dynamic system

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}),$$

where  $\mathbf{x}(t)$  is the measurement or output vector,  $\mathbf{u}(t)$  is the known system input vector and  $\boldsymbol{\theta}$  is the parameter vector, it is necessary to verify whether each unique system output corresponds to a unique set of parameter values, that is,

$$\mathbf{x}(\mathbf{u}, \boldsymbol{\theta}_1) = \mathbf{x}(\mathbf{u}, \boldsymbol{\theta}_2) \Rightarrow \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$$

is satisfied either globally or locally if the model is identifiable. Note that usually  $\mathbf{f} = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta})$  does not explicitly depend on  $t$  and  $\mathbf{u}(t)$ , therefore, it is equivalent to justify whether

$$\mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta}_1) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta}_2) \Rightarrow \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2 \quad (4.4)$$

holds globally or locally.

To test the identifiability of the parameters originated from the payoff matrix in general replicator dynamics, we use the theorem of equivalence between the replicator equations and the Lotka-Volterra

equations. Two systems of differential equations are called equivalent if one can be transformed into another by a certain change of variables. In particular the change of variable transforms solutions of one set of equations into solutions of another. In general, the problem of equivalence between differential equations consists in determining whether two equation systems are equivalent up to a given class of transformations. The theorem of equivalence between them is

**Theorem 4.3.1.** *The general Lotka-Volterra equations describe the interaction of  $n - 1$  species with  $n \geq 3$  and are of the form*

$$\dot{y}_i = y_i(r_i + \sum_{j=1}^{n-1} b_{ij}y_j) \quad i = 1, \dots, n-1, \quad (4.5)$$

where  $y_i \geq 0$  is the abundance or count of species  $i$ ,  $r_i$  is the growth rate of species  $i$ , and the interaction between species  $i$  and  $j$  is given by  $b_{ij}$ . The general replicator equations of  $n$  strategies are

$$\dot{x}_i = x_i \left[ \sum_{j=1}^n a_{ij}x_j - \sum_{j=1}^n \sum_{k=1}^n a_{jk}x_jx_k \right] \quad i = 1, \dots, n. \quad (4.6)$$

Then the Lotka-Volterra equations (4.5) with the parameters  $r_i = a_{in} - a_{nn}$  and  $b_{ij} = a_{ij} - a_{nj}$  are equivalent to the replicator equations (4.6).

The equivalence can be shown with the transformation  $y = \sum_{i=1}^{n-1} y_i$ ,  $x_i = y_i/(1+y)$  for  $i = 1, \dots, n-1$  and  $x_n = 1/(1+y)$ . The detailed proof is provided in Appendix A. From the equivalence of the two equations, we see that the structural identifiability of the parameters  $\{a_{ij} : i, j = 1, \dots, n\}$  in the replicator equations are closely related with that of the parameters  $\{r_i, b_{ij} : i, j = 1, \dots, n-1\}$  in the Lotka-Volterra equations.

To study the structural identifiability of the parameters in the Lotka-Volterra equations, we implement the direct test in (4.4). Let  $\theta = (r_1, \dots, r_{n-1}, b_{11}, \dots, b_{n-1, n-1})$ ,  $f_i(\mathbf{y}, \theta) = y_i(r_i + \sum_{j=1}^{n-1} b_{ij}y_j)$ ,  $i = 1, \dots, n-1$  and  $\mathbf{f}(\mathbf{y}, \theta) = (f_1(\mathbf{y}, \theta), \dots, f_{n-1}(\mathbf{y}, \theta))$ . From  $\mathbf{f}(\mathbf{y}, \theta) = \mathbf{f}(\mathbf{y}, \theta')$ , we have

$$y_i(r_i + \sum_{j=1}^{n-1} b_{ij}y_j) = y_i(r'_i + \sum_{j=1}^{n-1} b'_{ij}y_j), \quad i = 1, \dots, n-1. \quad (4.7)$$

Solving the above  $n-1$  equations, we obtain

$$r_i = r'_i, \quad i = 1, 2, \dots, n-1;$$

$$b_{ij} = b'_{ij}, \quad i, j = 1, 2, \dots, n-1,$$

which indicates that the parameters  $\theta = (r_1, \dots, r_{n-1}, b_{11}, \dots, b_{n-1, n-1})$  are identifiable. Thus we have the following identifiability theorem

**Theorem 4.3.2.** *For the general Lotka-Volterra equations*

$$\dot{y}_i = y_i(r_i + \sum_{j=1}^{n-1} b_{ij}y_j) \quad i = 1, \dots, n-1,$$

*the parameters  $(r_1, \dots, r_{n-1}, b_{11}, \dots, b_{n-1, n-1})$  are identifiable. From the equivalence of the Lotka-Volterra equations and the replicator equations of  $n$  strategies of the form*

$$\dot{x}_i = x_i \left[ \sum_{j=1}^n a_{ij}x_j - \sum_{j=1}^n \sum_{k=1}^n a_{jk}x_jx_k \right] \quad i = 1, \dots, n,$$

*the parameters  $\{a_{ij}, i, j = 1, \dots, n\}$  are not identifiable. However,  $n(n-1)$  combinations of the parameters  $\{a_{in} - a_{nn}, i = 1, \dots, n-1; a_{ij} - a_{nj}, i, j = 1, \dots, n-1\}$  are identifiable.*

In real life applications to these models, initial conditions may not be known. If the initial conditions of the frequencies for all species are considered to be parameters instead of known constants, then the conclusions on the identifiability of the parameters in the payoff matrix will still hold. There is not any literature dealing with the identifiability of the initial conditions in differential equations, especially the Lotka-Volterra type equations and the replicator equations. The identifiability of the initial conditions are defined in a way which is the same as the parameters in the dynamics. Thus, we conclude that the initial value parameters are uniquely determined given the output  $\mathbf{y}(t, \theta, \mathbf{y}_0)$  at initial time  $t_0$  since  $\mathbf{y}_0 = \mathbf{y}(t_0, \theta, \mathbf{y}_0)$  and if

$$\mathbf{y}(t_0, \theta_1, \mathbf{y}_{0,1}) = \mathbf{y}(t_0, \theta_2, \mathbf{y}_{0,2}),$$

then

$$\mathbf{y}_{0,1} = \mathbf{y}(t_0, \theta_1, \mathbf{y}_{0,1}) = \mathbf{y}(t_0, \theta_2, \mathbf{y}_{0,2}) = \mathbf{y}_{0,2},$$

which indicates that  $\mathbf{y}_0$  is identifiable. So from the perspective of mathematical identifiability, the initial state parameters  $\mathbf{y}_0 = \mathbf{y}(t_0, \theta, \mathbf{y}_0)$  are trivially identifiable as long as the states are observed at  $t_0$ , and they will not affect the identifiability of the parameters in the payoff matrix.

### 4.3.2 Dispersion Estimator

For the NB distribution, the variance is parameterized as

$$Var(Y_{ijm}) = \mu_{ijm} + \phi_{ij}\mu_{ijm}^2 \quad (4.8)$$

where  $\mu_{ijm}$  is the mean of the NB distribution and  $\phi_{ij}$  is the dispersion parameter. In the analysis of sequence data,  $\phi_{ij}$  accounts for the mouse-to-mouse variability. Robinson and Smyth (60) described several methods to estimate  $\phi_{ij}$ , and Nelder and Lee compared dispersion estimators and found that pseudo-likelihood methods are almost always less efficient than quasi-likelihood methods (57). In this article, we estimate  $\phi_{ij}$  by the quasi-likelihood method. As in (61) we estimate  $\phi_{ij}$  without imposing structure on  $\mu_{ijm}$ , and then condition on the estimates  $\hat{\phi}_{ij}$  in subsequent sections.

Suppose  $y_{ij1}, \dots, y_{ijM}$  are independent and  $\text{NB}(\mu_{ijm} = S_{jm}\lambda_{ij}, \phi_{ij})$  distributed, where  $S_{jm}$  is the total sample size and  $\lambda_{ij}$  represents the proportion of the population that is OTU  $i$  at time  $t_j$ . In this simple case, the MLE of  $\lambda_{ij}$  will depend on  $\phi_{ij}$  and maximum likelihood estimation of the two parameters  $(\lambda_{ij}, \phi_{ij})$  proceeds jointly for each  $i$  and  $j$ . We describe the quasi-likelihood approach for estimating NB dispersion based on  $\mu_{ijm} = S_{jm}\lambda_{ij}$ . The estimating equation is

$$2 \sum_{m=1}^M \left\{ y_{ijm} \log \left[ \frac{y_{ijm}}{\hat{\mu}_{ijm}} \right] - (y_{ijm} + \phi_{ij}^{-1}) \log \left[ \frac{y_{ijm} + \phi_{ij}^{-1}}{\hat{\mu}_{ijm} + \phi_{ij}^{-1}} \right] \right\} = M - 1, \quad \forall i, j, \quad (4.9)$$

where  $\hat{\mu}_{ijm} = S_{jm}\hat{\lambda}_{ij}$  and  $\hat{\lambda}_{ij}$  is the MLE of  $\lambda_{ij}$  given  $\phi_{ij}$ . After estimating  $\hat{\phi}_{ij}$  from equation (4.9), the only remaining unknown parameters are those connect with the mean component in the NB mixture model considered next.

### 4.3.3 NB Mixture Model

As suggested from the similarities observed in the dynamic profiles of OTUs in Figure 4.1, we will use soft clustering to group the OTUs into  $K$  groups. Suppose the  $n$  OTUs are independently assigned to these groups with the same prior probability. Let  $Z_i$  be the latent random variable indicating the group of OTU  $i$  and  $P(Z_i = k) = p_k, k = 1, \dots, K, \forall i$ . We assume that OTUs assigned to the same group follow identical dynamics. Thus, their relative proportions within the group do not change over time. In addition, because we observe sequences from OTUs, rather than the OTUs themselves, OTU  $i$  contributes a fixed amount  $d_{ik}$  to the sequence data originating from group  $k$ . Given the group assignment, the counts of sequences from OTU  $i$  at time  $t_j$  in mouse  $m$  are assumed to be independent NB, specifically

$$y_{ijm}|Z_i = k \sim \text{NB}(s_{jm}d_{ik}f_{kj}, \hat{\phi}_{ij}) \quad i = 1, \dots, n, \quad (4.10)$$

where  $k \in \{1, \dots, K\}$  is the group number,  $p_k d_{ik}$  is the relative abundance of sequences from OTU  $i$  in group  $k$  with constraint  $p_k \sum_{i=1}^n d_{ik} = 1$ ,  $f_{kj}$  is the proportion of group  $k$  at time  $t_j$ , and  $s_{jm}$  is the OTU sample size at time  $t_j$  in mouse  $m$ . The assumption  $p_k \sum_{i=1}^n d_{ik} = 1$  for all  $k$  yields

$$\begin{aligned} E \left( \sum_{i=1}^n y_{ijm} \right) &= \sum_{i=1}^n \sum_{k=1}^K p_k s_{jm} d_{ik} f_{kj} \\ &= s_{jm} \sum_{k=1}^K f_{kj} \left( p_k \sum_{i=1}^n d_{ik} \right) \\ &= s_{jm} \sum_{k=1}^K f_{kj} \cdot 1 \\ &= s_{jm}, \end{aligned}$$

which suggests that  $S_{jm} = \sum_{i=1}^n y_{ijm}$  is an unbiased estimator of  $s_{jm}$ . Also let  $\mathbf{y}_{jm} = (y_{1jm}, y_{2jm}, \dots, y_{njm})$  be the OTU counts at time  $t_j$  in mouse  $m$ , then the distribution of  $\mathbf{y}_{jm}$  for  $j = 1, \dots, N, m = 1, \dots, M$  conditioned on  $Z_1, \dots, Z_n$  is

$$y_{1jm}, \dots, y_{njm} | Z_1 = k_1, \dots, Z_n = k_n \sim \text{NB}(s_{jm} d_{1,k_1} f_{k_1,j}, \hat{\phi}_{1j}) \cdot \dots \cdot \text{NB}(s_{jm} d_{n,k_n} f_{k_n,j}, \hat{\phi}_{nj}).$$

To describe the dynamics of the proportions or frequencies of the  $K$  groups, we assume replicator dynamics

$$\dot{f}_i = f_i \left[ \sum_{j=1}^K a_{ij} f_j - \sum_{j=1}^K \sum_{k=1}^K a_{jk} f_j f_k \right] \quad i = 1, \dots, K,$$

where  $f_i(t), i = 1, \dots, K$  is the frequency of group  $i$  at time  $t$ , and  $a_{ij}$ 's are entries of the payoff matrix. Since the payoff matrix  $[a_{ij}]$  is not identifiable as proved in section 4.3.1, we reparameterize in terms of  $K$  by  $K$  matrix  $[b_{ij}]_{K \times K}$  with the  $K$  elements in last row set to be zero. The corresponding replicator equations are

$$\dot{f}_i = f_i \left[ \sum_{j=1}^K b_{ij} f_j - \sum_{j=1}^{K-1} \sum_{k=1}^K b_{jk} f_j f_k \right] \quad i = 1, \dots, K-1, \quad (4.11)$$

with  $f_K = 1 - \sum_{i=1}^{K-1} f_i$ . Denote  $\mathbf{b} = (b_{11}, \dots, b_{K(K-1)})$ ,  $\mathbf{s} = (s_{11}, \dots, s_{NM})$ ,  $\mathbf{d} = (d_{11}, \dots, d_{nK})$

and  $\underline{\mathbf{p}} = (p_1, \dots, p_K)$ , then the joint distribution of  $\underline{\mathbf{y}} = (\mathbf{y}_{11}, \dots, \mathbf{y}_{NM})$  and  $\underline{\mathbf{Z}} = (Z_1, \dots, Z_n)$  is

$$\begin{aligned} L(\underline{\mathbf{y}}, \underline{\mathbf{Z}}, \underline{\boldsymbol{\theta}}) &= P(\underline{\mathbf{y}}, \underline{\mathbf{Z}} | \underline{\boldsymbol{\theta}}) \\ &= \left[ \prod_{j=1}^N \prod_{m=1}^M P(\mathbf{y}_{jm} | \underline{\mathbf{Z}}, \underline{\boldsymbol{\theta}}) \right] P(\underline{\mathbf{Z}}) \\ &= \prod_{k_1=1}^K \dots \prod_{k_n=1}^K \prod_{i=1}^n \left[ p_{k_i} \prod_{j=1}^N \prod_{m=1}^M \frac{\Gamma(y_{ijm} + \hat{\phi}_{ij}^{-1})}{\Gamma(\hat{\phi}_{ij}^{-1})\Gamma(y_{ijm} + 1)} \left( \frac{1}{1 + s_{jm}d_{i,k_i}f_{k_i,j}\hat{\phi}_{ij}} \right)^{\hat{\phi}_{ij}^{-1}} \right. \\ &\quad \left. \left( \frac{s_{jm}d_{i,k_i}f_{k_i,j}}{\hat{\phi}_{ij}^{-1} + s_{jm}d_{i,k_i}f_{k_i,j}} \right)^{y_{ijm}} \right]^{I[Z_i=k_i]}, \end{aligned}$$

where  $\underline{\boldsymbol{\theta}} = (\underline{\mathbf{p}}, \underline{\mathbf{b}}, \underline{\mathbf{s}}, \underline{\mathbf{d}})$  and  $f_{k_i,j}$  is the frequencies of group  $k_i$  in replicator dynamics evaluated at  $t_j$ .

Then the complete data log-likelihood function is

$$\begin{aligned} l(\underline{\mathbf{y}}, \underline{\mathbf{Z}}, \underline{\boldsymbol{\theta}}) &= \log L(\underline{\mathbf{y}}, \underline{\mathbf{Z}}, \underline{\boldsymbol{\theta}}) \\ &= \sum_{k_1=1}^K \dots \sum_{k_n=1}^K I[Z_1 = k_1] \dots I[Z_n = k_n] \left\{ (\log p_{k_1} + \dots + \log p_{k_n}) + \sum_{i=1}^n \sum_{j=1}^N \sum_{m=1}^M [\log \Gamma(y_{ijm} + \right. \\ &\quad \left. \hat{\phi}_{ij}^{-1}) - \log \Gamma(\hat{\phi}_{ij}^{-1}) - \Gamma(y_{ijm} + 1) - (\hat{\phi}_{ij}^{-1} + y_{ijm}) \log(1 + s_{jm}d_{i,k_i}f_{k_i,j}\hat{\phi}_{ij}) + y_{ijm} \log(s_{jm} \cdot \right. \\ &\quad \left. d_{i,k_i}f_{k_i,j}\hat{\phi}_{ij})] \right\}. \end{aligned}$$

The parameters  $\underline{\mathbf{b}}$  are the key parameters we would like to estimate. It is computationally intensive to directly integrate the replicator equations (4.11) to estimate the parameters. We use a strategy originally introduced by Ramsay et. al (2007), which uses smoothing splines  $\mathbf{w}(\cdot)$  to approximate solution curves  $\mathbf{f}(\cdot)$ . Then the log-likelihood  $l(\underline{\mathbf{y}}, \underline{\mathbf{Z}}, \underline{\boldsymbol{\theta}})$  can be approximated by

$$\begin{aligned} l'(\underline{\mathbf{y}}, \underline{\mathbf{Z}}, \underline{\boldsymbol{\theta}}^\Delta) &= \sum_{k_1=1}^K \dots \sum_{k_n=1}^K I[Z_1 = k_1] \dots I[Z_n = k_n] \left\{ (\log p_{k_1} + \dots + \log p_{k_n}) + \sum_{i=1}^n \sum_{j=1}^N \sum_{m=1}^M [\log \Gamma(y_{ijm} + \right. \\ &\quad \left. \hat{\phi}_{ij}^{-1}) - \log \Gamma(\hat{\phi}_{ij}^{-1}) - \Gamma(y_{ijm} + 1) - (\hat{\phi}_{ij}^{-1} + y_{ijm}) \log(1 + s_{jm}d_{i,k_i}w_{k_i,j}(\underline{\mathbf{c}})\hat{\phi}_{ij}) + y_{ijm} \log(\right. \\ &\quad \left. s_{jm}d_{i,k_i}w_{k_i,j}(\underline{\mathbf{c}})\hat{\phi}_{ij})] \right\}, \end{aligned}$$

where  $\underline{\mathbf{c}}$  are the smoothing spline coefficients,  $\underline{\boldsymbol{\theta}}^\Delta = (\underline{\boldsymbol{\theta}}, \underline{\mathbf{c}})$  and  $w_{k_i,j}(\underline{\mathbf{c}}) = w_{k_i}(t_j | \underline{\mathbf{c}})$ . To control the smoothness of the splines and insure they are close to  $\mathbf{f}(\cdot)$ , we minimize a penalized negative log likelihood

$$l^\Delta(\underline{\mathbf{y}}, \underline{\mathbf{Z}}, \underline{\boldsymbol{\theta}}^\Delta | \lambda) = l'(\underline{\mathbf{y}}, \underline{\mathbf{Z}}, \underline{\boldsymbol{\theta}}^\Delta) + \lambda \text{PEN}(\mathbf{w} | \underline{\boldsymbol{\theta}}^\Delta),$$

where the penalty term  $\text{PEN}(\mathbf{w}|\boldsymbol{\theta}^\Delta)$  is

$$\text{PEN}(\mathbf{w}|\boldsymbol{\theta}^\Delta) = \sum_{k=1}^{K-1} \int_{[t_1, t_N]} \left[ \frac{dw_k(t|\mathbf{c})}{dt} - g_k(\mathbf{w}(t|\mathbf{c}), \mathbf{b}) \right]^2 dt,$$

and  $\{g_k(\cdot)\}_{k=1}^{K-1}$  are the right hand side terms in equations (4.11).

#### 4.3.4 EM Algorithm

The EM algorithm can be applied to find the MLE of the observed data log likelihood  $E_{\mathbf{Z}|\mathbf{y}} \left[ l^\Delta(\mathbf{y}, \mathbf{Z}, \boldsymbol{\theta}^\Delta | \lambda) \right]$  by iteratively applying the two-steps:

**Expectation step (E step)** At iteration  $t$ , calculate the expected value of the negative log-likelihood function with respect to the conditional distribution of  $\mathbf{Z}$  given  $\mathbf{y}$  under the current estimate  $\boldsymbol{\theta}^{\Delta(t)}$  of the parameters.

$$\begin{aligned} Q(\boldsymbol{\theta}^\Delta | \boldsymbol{\theta}^{\Delta(t)}) &= E_{\mathbf{Z}|\mathbf{y}} \left[ l^\Delta(\mathbf{y}, \mathbf{Z}, \boldsymbol{\theta}^\Delta | \lambda) \right] \\ &= - \sum_{k_1=1}^K \cdots \sum_{k_n=1}^K p_{1,k_1}^{(t)} \cdots p_{n,k_n}^{(t)} \left\{ (\log p_{k_1} + \dots + \log p_{k_n}) + \sum_{i=1}^n \sum_{j=1}^N \sum_{m=1}^M [\log \Gamma(y_{ijm} + \hat{\phi}_{ij}^{-1}) - \right. \\ &\quad \left. \log \Gamma(\hat{\phi}_{ij}^{-1}) - \Gamma(y_{ijm} + 1) - (\hat{\phi}_{ij}^{-1} + y_{ijm}) \log(1 + s_{jm} d_{i,k_i} w_{k_i,j}(\mathbf{c}) \hat{\phi}_{ij}) + y_{ijm} \log(s_{jm} d_{i,k_i} \cdot \right. \\ &\quad \left. w_{k_i,j}(\mathbf{c}) \hat{\phi}_{ij})] \right\} + \lambda \sum_{k=1}^{K-1} \int_{[t_1, t_N]} \left[ \frac{dw_k(t|\mathbf{c})}{dt} - g_k(\mathbf{w}(t|\mathbf{c}), \mathbf{b}) \right]^2 dt, \end{aligned}$$

$$\text{where } p_{i,k_i}^{(t)} = P(Z_i = k_i | \mathbf{y}_i, \boldsymbol{\theta}^{\Delta(t)}) = \frac{P(Z_i = k_i, \mathbf{y}_i, \boldsymbol{\theta}^{\Delta(t)})}{P(\mathbf{y}_i, \boldsymbol{\theta}^{\Delta(t)})} \approx \frac{p_{k_i}^{(t)} \prod_{j=1}^N \prod_{m=1}^M h(y_{ijm} | s_{jm}^{(t)} d_{i,k_i}^{(t)} w_{k_i,j}, \hat{\phi}_{ij})}{\sum_{k_i=1}^K p_{k_i}^{(t)} \prod_{j=1}^N \prod_{m=1}^M h(y_{ijm} | s_{jm}^{(t)} d_{i,k_i}^{(t)} w_{k_i,j}, \hat{\phi}_{ij})} \text{ with}$$

$h(\cdot | \mu, \hat{\phi})$  being the probability density function of NB distribution. The conditional distribution is approximate because  $f_k(\cdot)$  is substituted with  $w_k(\cdot | \mathbf{c})$  for all  $k = 1, \dots, K$ .

**Minimization step (M step)** Find the parameter  $\boldsymbol{\theta}^\Delta = (\mathbf{p}, \mathbf{b}, \mathbf{s}, \mathbf{d}, \mathbf{c})$  that minimizes the quantity

$$\boldsymbol{\theta}^{\Delta(t+1)} = \arg \min_{\boldsymbol{\theta}^\Delta} Q(\boldsymbol{\theta}^\Delta | \boldsymbol{\theta}^{\Delta(t)}).$$

We perform the minimization in blocks

$$(m1) \quad \underline{s}^{(t+1)} = \arg \min_{\underline{s}} Q((\underline{p}^{(t)}, \underline{s}, \underline{d}^{(t)}, \underline{b}^{(t)}, \underline{c}^{(t)}) | \underline{\theta}^{\Delta(t)})$$

$$(m2) \quad (\underline{p}^{(t+1)}, \underline{d}^{(t+1)}) = \arg \min_{\underline{p}, \underline{d}} Q((\underline{p}, \underline{s}^{(t+1)}, \underline{d}, \underline{b}^{(t)}, \underline{c}^{(t)}) | \underline{\theta}^{\Delta(t)})$$

$$(m3) \quad (\underline{b}^{(t+1)}, \underline{c}^{(t+1)}) = \arg \min_{\underline{b}, \underline{c}} Q((\underline{p}^{(t+1)}, \underline{s}^{(t+1)}, \underline{d}^{(t+1)}, \underline{b}, \underline{c}) | \underline{\theta}^{\Delta(t)}).$$

One full cycle (m1), (m2) and (m3) results in updates  $\underline{\theta}^{\Delta(t)}$  to  $\underline{\theta}^{\Delta(t+1)}$ .

For the first optimization problem, the explicit solution of  $\underline{s}^{(t+1)}$  can not be derived and it is necessary to implement the quasi-newton optimization algorithm to solve

$$\min_{\underline{s}} Q((\underline{p}^{(t)}, \underline{s}, \underline{d}^{(t)}, \underline{b}^{(t)}, \underline{c}^{(t)}) | \underline{\theta}^{\Delta(t)}). \quad (4.12)$$

In the second problem, there are no closed form solutions for both  $\underline{p}^{(t+1)}$  and  $\underline{d}^{(t+1)}$ . The optimization is based on interior point numerical algorithm for the constraint optimization problem

$$\min_{\underline{p}, \underline{d}} Q((\underline{p}, \underline{s}^{(t+1)}, \underline{d}, \underline{b}^{(t)}, \underline{c}^{(t)}) | \underline{\theta}^{\Delta(t)})$$

$$\text{subject to} \quad \underline{p} \geq 0, \underline{d} \geq 0, \quad (4.13)$$

$$p_k \sum_{i=1}^n d_{ik} = 1, \quad i = 1, \dots, K.$$

And the third problem

$$\min_{\underline{b}, \underline{c}} Q((\underline{p}^{(t+1)}, \underline{s}^{(t+1)}, \underline{d}^{(t+1)}, \underline{b}, \underline{c}) | \underline{\theta}^{\Delta(t)}) \quad (4.14)$$

is tackled by using the inner-outer optimization procedure of (59) which is an iterative inner-outer quasi-Newton method (BFGS) based optimization algorithm. Details of the optimizations (m1), (m2) and (m3) are provided in the Appendix C.

#### 4.3.5 Initialization

The EM algorithm has the guaranteed descent property, but it may converge to a local minimum. It is well known that initial parameter values can impact both the speed of convergence and minimum

achieved. To propose good initial values, we apply k-means to principal components to assign OTUs to initial clusters. Specifically, we take time series frequencies for each OTU, standardize them by centering and then normalizing, and merge the  $m$  replicates by concatenating them. We then apply principal component analysis (PCA) (50), retain the top two components and cluster the OTUs using k-means. The initial values of  $\tilde{s}^{(0)}$  are chosen to be the observed sample sizes  $S_{jm}$ . From the initial clustered groups, the initial values of  $\tilde{p}^{(0)}$  are estimated to be the proportions of OTUs within each group. And initial guess of  $d_{ik}^{(0)}$  can be derived from the mean component of (4.10) when OTU  $i$  is assigned to intial group  $k$ , while initial guesses of  $d_{sk}^{(0)}$  for the OTU  $s$  not assigning to group  $k$  are set equal while satisfying the constraints given  $\tilde{p}^{(0)}$ . The initial values of  $\tilde{b}^{(0)}$  are derived based on least squares fitting of equations (4.11) to the frequencies of initial  $K$  groups. The initial guesses of the smoothing spline parameters  $\tilde{c}^{(0)}$  are computed via the interpolation of logit transformed splines to the initial group frequencies curves.

## 4.4 Simulation Study

### 4.4.1 Data Simulation

We undertook a simulation study to test the feasibility and acceptability of the proposed estimation procedure. Note, there are 4 types of parameters  $\tilde{b}$ ,  $\tilde{d}$ ,  $\tilde{s}$ , and  $\tilde{p}$ . The control parameters for the simulation are chosen as  $K = 3$ ,  $n = 17$ ,  $N = 21$  and  $M = 5$ . The dispersion parameters  $\phi_{ij}$ 's are randomly sampled as  $\eta_i \text{Gamma}(0.5, 2)$ , where  $\eta_i$ 's are uniformly sampled in the closed interval  $[0.5, 2]$ . We set  $\tilde{b} = (-1, 2, -1, -2, 1, 1)$ ,  $\tilde{p} = (0.3, 0.5, 0.2)$ , and sample  $s_{jm}$ 's as integers between 2000 and 3000. For each  $k$ , the  $d'_{ik}$ 's are randomly sampled from an unit simplex  $\{d'_{ik} : 1 \leq i \leq n, d'_{ik} \geq 0, \sum_{i=1}^n d'_{ik} = 1\}$  and then  $d_{ik} = d'_{ik}/p_k$ , so  $d_{ik}$  satisfy  $p_k \sum_{i=1}^n d_{ik} = 1$ . Based on the chosen model parameters, we simulate the  $y_{ijm}$ 's independently from the NB mixture model. That is, first we generated the true frequency curves for the three groups as show in Figure 4.2(a), by numerically solving the replicator equations (Matlab function: ode45) at  $N = 21$  time points, and then randomly assigned the 17 OTUs to the three groups with the probability vector  $\tilde{p}$ . For OTU  $i$  in group  $k_i$ , we then sampled each  $y_{ijm}$  from a NB distribution with mean parameter  $s_{jm} d_{i,k_i} f_{k_i,j}$  and dispersion parameter  $\phi_{ij}$ . The simulated data for 17 OTUs observed at 21 equally spaced time points in 5 mice are shown in Figure 4.3.

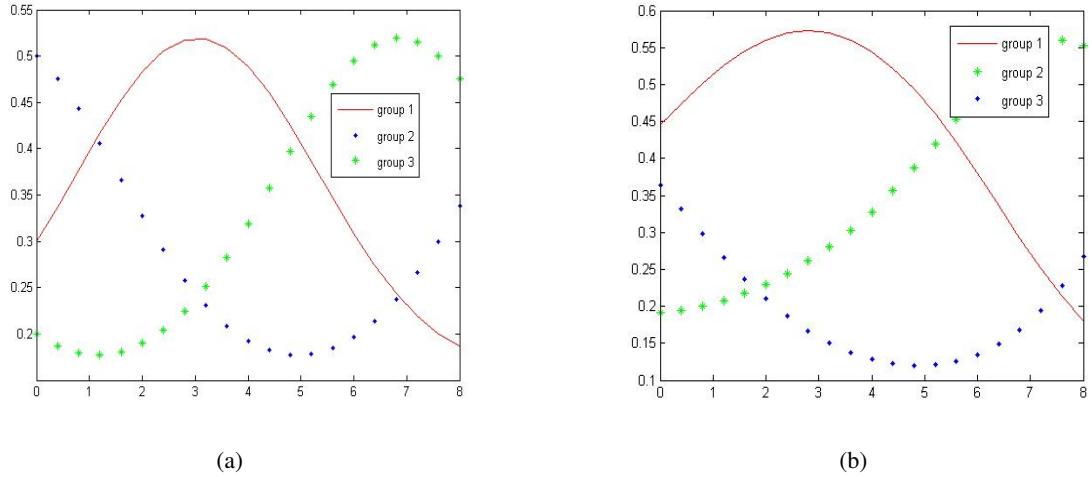


Figure 4.2 Frequencies of three groups solved by integrating the replicator equations at (a) the true parameters or (b) the parameters estimated during initialization.

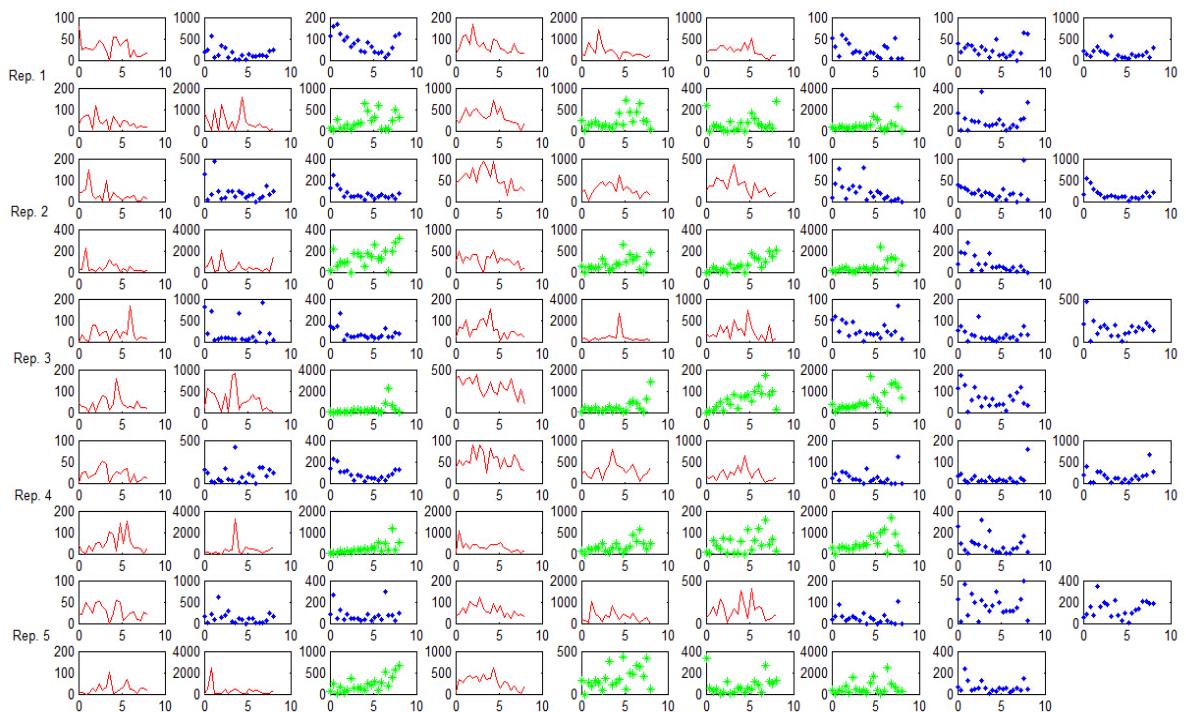


Figure 4.3 Simulated data for 17 OTUs of 5 mice at 21 time points.

#### 4.4.2 Results

We implemented our estimation procedure and applied it to the simulation data. All estimation is conditioned on the quasi-likelihood estimates at the dispersion parameters which are found by solving  $nN$  nonlinear estimating equations (4.9). A summary of the estimated dispersion parameters is provided in Figure 4.4. Most of the dispersion parameters are well estimated; only a few of the large magnitude dispersion parameters are poorly estimated. Large dispersion parameters tend to be overestimated when the number of replicates is small (45).

Table 4.1 Initial values of  $\underline{b}^{(0)}$ .

|                | $b_{11}$ | $b_{12}$ | $b_{13}$ | $b_{21}$ | $b_{22}$ | $b_{23}$ |
|----------------|----------|----------|----------|----------|----------|----------|
| Initial values | -0.1035  | -1.1434  | 1.1000   | -1.4024  | 1.0779   | 0.4471   |

We next compute initial guesses of all parameters. Our initialization scheme estimated payoff matrix parameters  $\underline{b}^{(0)}$  as shown in Table 4.1, which produce frequency  $f_k(\cdot)$  curves similar to the truth (Figure 4.2).

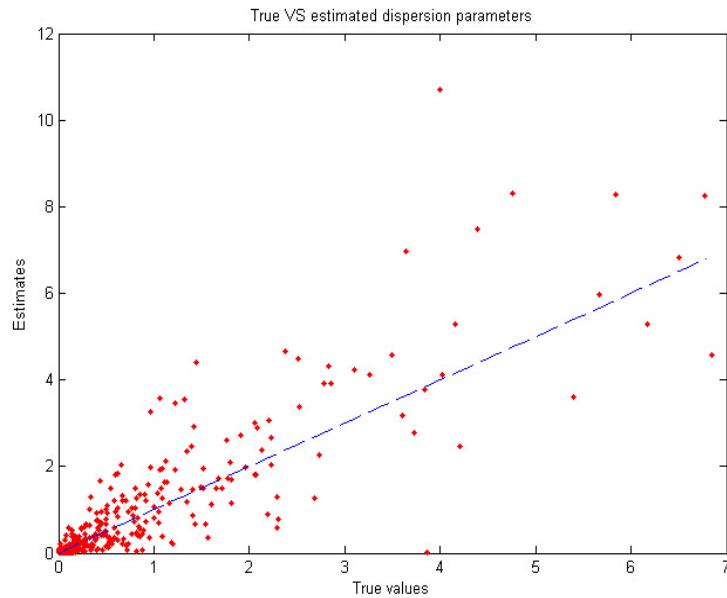


Figure 4.4 Summary of estimated dispersion parameters.

The EM algorithm converged in five iterations, using the stop criterion that sum of relative changes over the four types of parameters is no more than  $10^{-4}$ . A summary of the estimated  $s_{jm}$ 's and  $d_{ik}$ 's

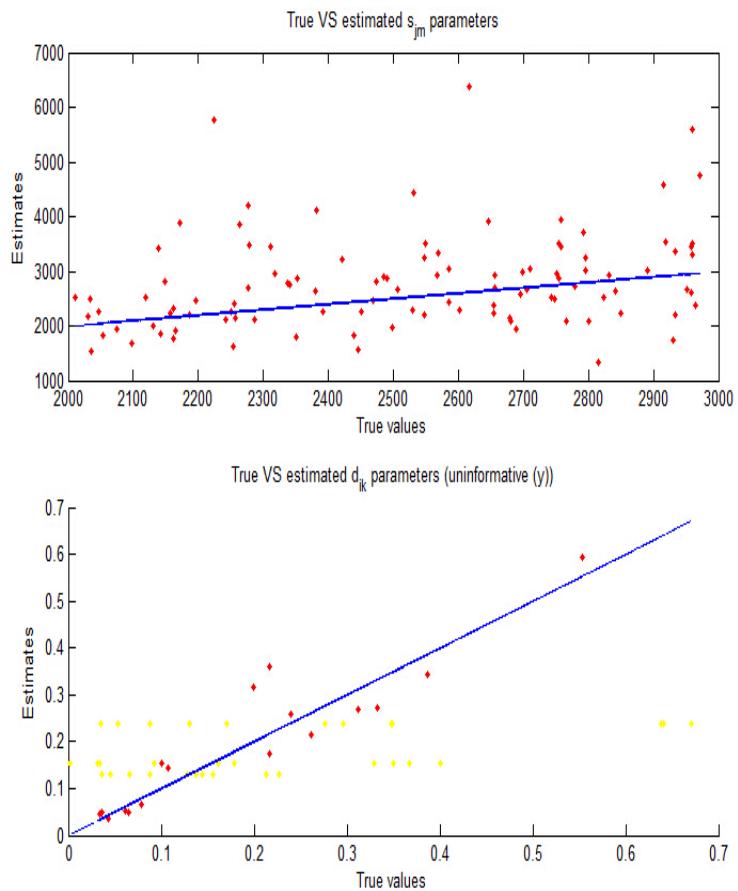


Figure 4.5 Summary of estimates for parameters  $s_{jm}$  and  $d_{ik}$ .

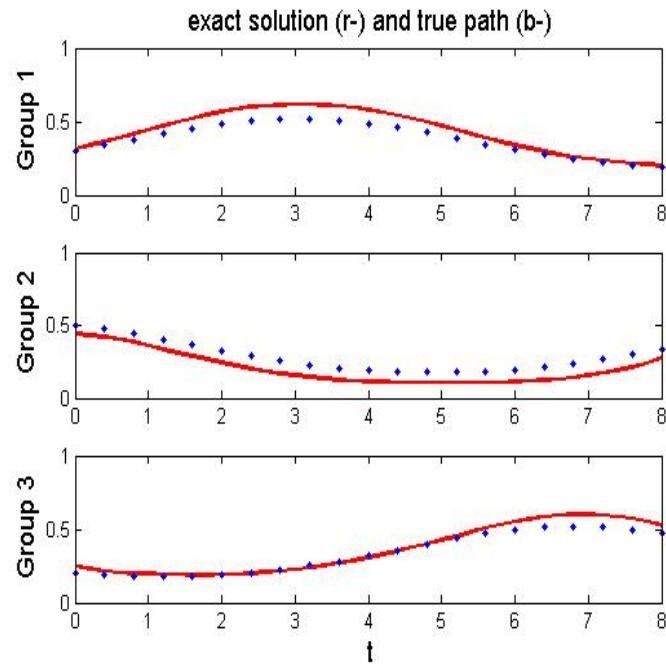


Figure 4.6 Comparison of fitted exact solution with solution of differential equations.

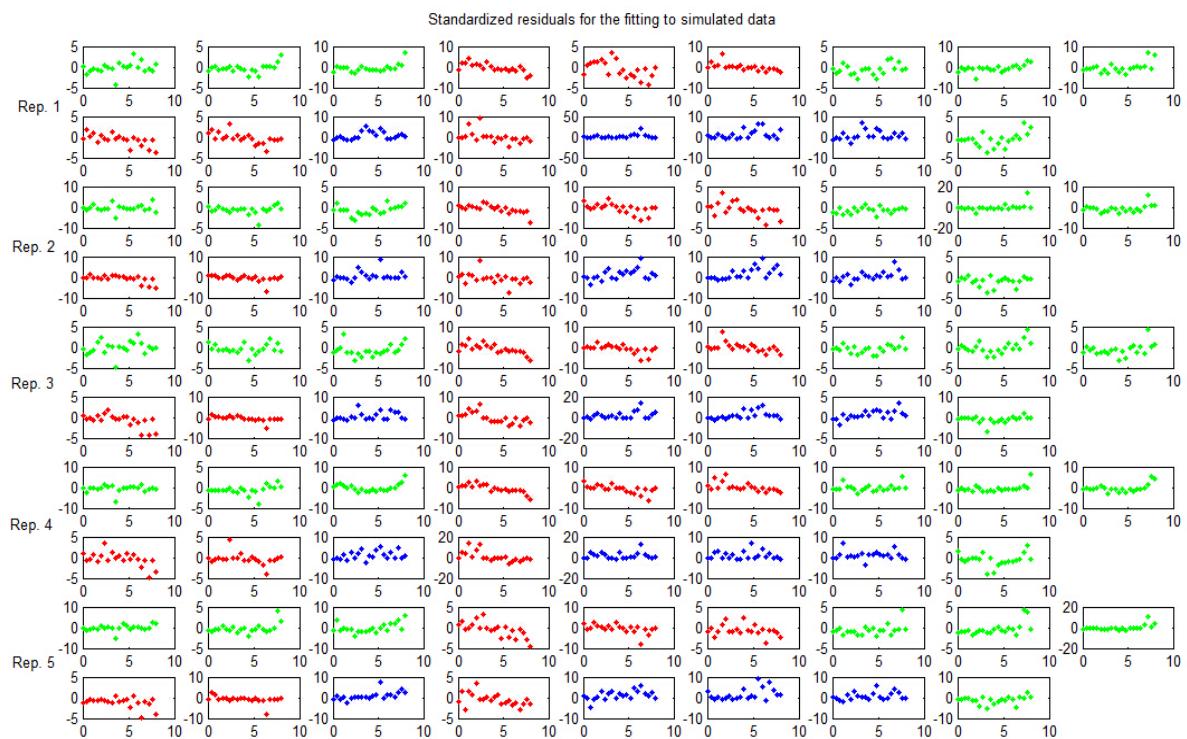


Figure 4.7 Standardized residuals for the fitting to the simulated data.

are plotted in Figure 4.5. The first plot in the figure suggests that a few  $s_{jm}$  are mis-estimated, but most are reasonably close to the true values. When OTU  $i$  has high posterior probability of being in group  $k$ , then there is no information to estimate  $d_{il}$  for  $l \neq k$ . These are shown yellow in scatter plot 4.5. The remaining  $d_{ik}$  are well-estimated. The true and estimated parameters in  $\underline{p}$  and  $\underline{b}$  are provided in Table 4.2. Parameter  $\underline{p}$  is well estimated considering it is based on a sample of only 17 OTUs. For the estimates of  $\underline{b}$ , we can see that four pairs of interactions between the three groups are estimated correctly in terms of signs of the interactions while the last two pairs are not. It is possible that the noisy data error, the smoothing spline approach, or poor initialization resulted in a local minimum and produced suboptimal estimates of the payoff parameters. There could also be some lack of sensitivity, or statistical identifiability issues in this simulation. Further investigation is needed to assess the cause. The temporal dynamics of the frequencies of three groups based the final estimates of the parameters are shown in Figure 4.6, and from the comparison of the exact frequencies curves and true frequencies paths for the three groups in the plot, we conclude that the estimated frequencies curves fit the true paths very well.

Table 4.2 True values and estimates of  $\underline{p}$  and  $\underline{b}$ .

|           | $p_1$  | $p_2$  | $p_3$  | $b_{11}$ | $b_{12}$ | $b_{13}$ | $b_{21}$ | $b_{22}$ | $b_{23}$ |
|-----------|--------|--------|--------|----------|----------|----------|----------|----------|----------|
| True      | 0.3    | 0.5    | 0.2    | -1       | 2        | -1       | -2       | 1        | 1        |
| Estimates | 0.3530 | 0.4119 | 0.2351 | -0.6494  | -1.1295  | 2.9130   | -3.0224  | 2.4077   | 1.5635   |

After fitting the NB mixture model to data, it is important to perform model diagnostics to assess the goodness of fit and validity of model assumptions. In simulation, we know model assumptions are satisfied, but the residual plots in Figure 4.7 reveal a few OTUs where the residuals across mice still show temporal patterns. This dependence in residuals is likely caused by disagreements between the replicator frequencies  $f_k(\cdot)$  and the smoothing spline approximats  $w_k(\cdot|\mathcal{C})$  because the data were simulated independently, conditional on the modeled means.

## 4.5 Discussion and Future Work

The estimation procedure is so far quite slow, so we were only able to perform one complete EM run. Thus it was impossible to assess what might have failed or what could be improved in the method.

For the future, it will be helpful to rerun the estimation while holding some subsets of the parameters at their true values. In this way, we can assess whether the observed misestimations are significantly impacting the estimation of  $\hat{b}$ , our main parameters of interest. In addition, we chose to use a smoothing spline approximation with 164 parameters. We should rerun the analysis with more or fewer smoothing spline parameters.

We used simulated data to assess our method for estimating interaction effects between species in mixed populations. We assumed the species interacted through replicator dynamics that hypothesize a specific payoff for interactions between pairs of species. Cooperating species  $i$  and  $j$  will receive positive payoffs,  $a_{ij}, a_{ji} > 0$ , from their interaction, while a parasitic species  $i$  will negatively impact its victim  $j$ ,  $a_{ji} < 0 < a_{ij}$ . Other interactions are also possible. In modern metagenomics datasets, there are many potentially interacting species, which leads to an explosion in the number of parameters. However, it is often true that many distinguishable species will interact similarly, so it is plausible to cluster species in homogeneously reacting groups. Our proposed model allows such grouping. Note, our grouping is partly genetic (reads are clustered into OTUs) and partly phenotypic (based on temporal dynamics). We feel this approach is better than pure genetic grouping, which treats species as less behaviorally plastic than they are. Future extensions could cluster genomes based on both criteria.

It is difficult to assess the performance of our method as the dataset size increases. Unfortunately, adding new data in any direction, more OTUs, more time points, or more mice, leads to concomitant increases in the number of parameters. The only way to increase the amount of data without adding more parameters is to increase  $s_{jm}$ . Of course, this also increases the variance. In real data applications, it may be possible to regularize  $s_{jm}$  across times or mice to control the number of parameters. Multiplex sequences may also help, and it may also work reasonably well to set  $s_{jm} = S_{jm}$ .

Another completely different approach is to dispense with the smoothing spline approximation. Instead, we can utilize a Newton-Raphson optimization, which should converge in fewer iterations, thus requiring relatively few numerical integrations of the replication equations and similar equations for the Hessians. Fortunately, for the cluster model, the payoff matrix is relatively small.

## CHAPTER 5. MARKOV RANDOM FIELD MODEL FOR ANALYSIS OF GAME DYNAMICS

### **Abstract**

Inverse problem is a well-studied mathematical problem and arise in many branches of science and mathematics. The conceptual understanding is that the transformation from data to model parameters is a result of the interaction of a physical system with the object that we wish to infer about. To estimate the parameters of the payoff matrix for linear game model, we introduce a statistical method—Markov Random Field Model which mainly deals with observations on a lattice using appropriate random variable distribution. The parameter estimation is a process of maximization of the Besag likelihood function over the parameter space and the inference of the parameters is based a MCMC method. Also we investigate the spatial effects of interaction region or neighborhood size on the estimated parameters for given data and analyze the nature of the game that we are interested in.

### **5.1 Introduction, Notations and Preliminaries**

In this section, we develop a spatial-temporal statistical model for symmetric  $2 \times 2$  games, where the payoff for each player is given by the payoff matrix

$$\left( \begin{array}{c|cc} & C & D \\ \hline C & R & S \\ D & T & P \end{array} \right) \quad (5.1)$$

where  $C$  and  $D$  stand for the population of cooperators and defectors, respectively. Consider the game defined on the lattice  $\Omega = \{1, 2, \dots, n\}^2 \subset \mathbb{Z}^2$ . The underlying structure of the game dynamics on  $[0, T] \times \Omega$  is decomposed into the temporal evolution, which is modeled via Markov chains, and the intrinsic spatial dependence, which is modeled by the Markov random fields. We focus on the

discrete time  $[0, T] := \{0, 1, \dots, T\}$  while the generalization to continuous time can be obtained by discretizations.

Denote  $s_i = (u, v)$  the point on  $\Omega$  with horizontal and vertical coordinates  $u, v \in \{1, 2, \dots, n\}$ . Let  $N = n^2$ . Consider  $\mathbf{Y}(t) = \{Y(s_i, t)\}_{i=1}^N$  to be the set of possible observations of a symmetric  $2 \times 2$  game on  $\Omega$  at time  $t$ , so that  $\mathbf{Y}(t)$  is a set of binary random variables on  $\{0, 1\}^{|\Omega|}$  and each  $Y(s_i, t)$  is a binary random variable on  $\{0, 1\}$ .

$N_i$  denotes the neighborhood of location  $s_i$ , i.e.  $N_i = \{s_j : s_j \text{ is in the neighbor of } s_i\}$ , which is assumed to be invariant with respect to time in our model. Naturally,  $Y(N_i, t) = \{Y(s_j, t)_{s_j \in N_i}\}$  denotes the collection of observations in the neighborhood of location  $s_i$  at time point  $t$ .

Furthermore, it is noticed that for the locations at which the focal player is surrounded by the same type of players, the probability of updating the central one with an alternative type is very small. The corresponding state of the process is "transiently forbidden", *i.e.* there exists state  $\omega \in \{0, 1\}^N$  such that

$$P(\mathbf{Y}_t = \omega \mid \mathbf{Y}_{t-1} \text{ has players surrounded by the same type}) \approx 0$$

which violates both the Besag positivity condition (66) and Kaiser-Cressie support conditions (75) for the modeling spatial structure with appropriate dependences. To circumvent this difficulty, we define the **set of "inactive" locations** by

$$\mathcal{I}_t = \left\{ (s_i, t) : \sum_{s_j \in N_i \cup \{s_i\}} y(s_j, t-1) = |N_i \cup \{s_i\}| \text{ or } 0 \right\} \quad (5.2)$$

for  $t = 0, 1, 2, \dots, T$  and the compliment  $\mathcal{I}_t^c = \Omega \setminus \mathcal{I}_t$  defines the **set of "active" locations**. Conditional on the history information of  $\mathbf{Y}_{t-1}$ , the sets of inactive locations and active locations are both random variables. The set of inactive locations is demonstrated in figure 5.1.

It is well known that we can reparameterize the game dynamics by population benefits during interactions to incorporate the Hamiltonian rule in the game model (67; 71). Furthermore, as discussed in (74) in terms of the equilibria, the game dynamics defined by the payoff matrix is equivalent to the corresponding replicator equations with parameters given by the payoff matrix. Hence, we have the following proposition claiming that the game is strategically invariant, by which we mean the equilibrium of evolution or optimal strategy is unchanged, subject to adding constants to each columns.

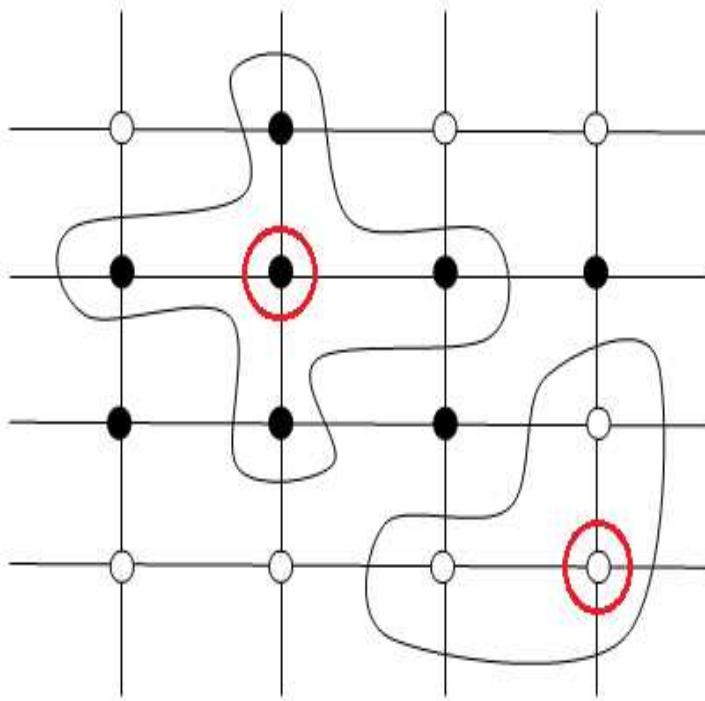


Figure 5.1 Illustration of set of inactive locations. The underlying geometry is 4 nearest neighborhood with fixed boundary. The highlighted two locations are inactive locations as they are surrounded by the same type of players/observations at time  $t$ . The evolution of observations on active locations may change the inactive location at  $t$  to active location at  $t + 1$ , which corresponds to the term "transient".

**Proposition 5.1.1.** *Consider the game that is equivalent to the replicator equation associated with the payoff matrix  $A$*

$$\frac{d\mathbf{x}}{dt} = \mathbf{x} \circ (A\mathbf{x}) - \mathbf{x}(\mathbf{x}' A \mathbf{x}) \quad (5.3)$$

*defined on simplex  $\mathcal{S}_n = \{\mathbf{x} \in \mathbb{R}_+^n \cup \{0\} \mid \mathbf{x} \cdot \mathbf{I} = 1\}$  where  $\mathbf{I} = (1, \dots, 1)'$ . The addition of a constant  $d_j$  to the  $j$ -th column of payoff matrix  $A$  does not change (5.3) on  $\mathcal{S}_n$ . Hence game defined by  $A$  is strategically invariant subject to adding constants to columns.*

*Proof.* W.L.O.G., adding  $d_j \in \mathbb{R}$  to the  $j$ -th column of  $A$ , so that  $\tilde{A} = (\mathbf{a}_1, \dots, \mathbf{a}_j + d_j \cdot \mathbf{1}, \dots, \mathbf{a}_n)$ .

The replicator dynamics (5.3) is therefore rewritten as, for  $i \in \{1, \dots, n\}$

$$\begin{aligned}
x'_i &= x_i \left( (\tilde{A}\mathbf{x})_i - \mathbf{x}' \tilde{A}\mathbf{x} \right) \\
\Rightarrow x'_i &= x_i \left\{ \sum_{k \neq j, k=1}^n a_{ik} x_k + (a_{ij} + d_j) x_j \right\} - x_i \left\{ \sum_{l=1}^n \sum_{k=1, k \neq j}^n a_{lk} x_l x_k + \sum_{l=1}^n (a_{lj} + d_j) x_l x_j \right\} \\
&= x_i \left( (A\mathbf{x})_i - \mathbf{x}' A\mathbf{x} \right) + d_j x_i x_j - d_j x_i x_j \left( \sum_{i=1}^n x_l \right) \\
&= x_i \left( (A\mathbf{x})_i - \mathbf{x}' A\mathbf{x} \right)
\end{aligned}$$

since  $\mathbf{x} \in \mathcal{S}_n$ . Hence, the replicator dynamics is invariant under translation by adding constants on each column.  $\square$

Therefore, the symmetric game defined by (5.1) is equivalent to the game defined by

$$\left( \begin{array}{c|cc} & C & D \\ \hline C & a & 0 \\ D & 0 & b \end{array} \right) \tag{5.4}$$

where  $a = R - T$  and  $b = P - S$ , and we can define the **local cooperator benefits** for location  $s_i$  at time  $t$  by

$$\pi_C^{i,t} = \frac{a}{|N_i \cup \{s_i\}|} \sum_{s_j \in N_i \cup \{s_i\}} (1 - y(s_j, t)) \tag{5.5}$$

and the **local defector benefits** for  $Y(s_i, t)$  by

$$\pi_D^{i,t} = \frac{b}{|N_i \cup \{s_i\}|} \sum_{s_j \in N_i \cup \{s_i\}} y(s_j, t). \tag{5.6}$$

## 5.2 Initial Configuration

The initial configuration  $\mathbf{Y}_0 = \{Y(s_i, 0)\}_{i=1}^N$  determines the following process and we define an auto-logistic model (66) as

$$P \{Y(s_i, 0) | Y(N_i, 0)\} = \frac{\exp[y(s_i, 0) A_i(y(N_i, 0))]}{1 + \exp[y(s_i, 0) A_i(y(N_i, 0))]} \tag{5.7}$$

where

$$A_i(y(N_i, 0)) = \log \left( \frac{\kappa_0}{1 - \kappa_0} \right) + \eta_0 \sum_{s_j \in N_i} (y(s_j, 0) - \kappa_0) \tag{5.8}$$

with  $\kappa_0 \in [0, 1]$ ,  $\eta_0 \geq 0$ . Hence, we assume that at the initial time point, the player's conditional state (either cooperator or defector) at a given location  $s_i$  depends only on the types of players of its neighborhood  $N_i$ . Notice that the first term in  $A_i(y(N_i, 0))$  is the natural parameter function for binary measure in the exponential form, and  $\kappa_0$ , unless the dependence structure  $\eta_0$  is strong, captures the marginal expectation of initial configuration.

Furthermore, (5.8) satisfies Besag positivity/symmetric condition (66; 72) so that the joint distribution for the initial configuration is well-defined, i.e.

$$P \{ \mathbf{Y}_0 = \mathbf{y}_0 \}$$

exists for all  $\mathbf{y}_0$  and let  $\Omega_0$  denote the collection of values after observations at  $t = 0$ ,  $\mathcal{C}_{\Omega_0}$  denote the set of possible configuration for  $\Omega_0$ , there is

$$\sum_{\mathcal{C}_{\Omega_0}} P(\Omega_0) = 1. \quad (5.9)$$

### 5.3 Development of Temporal Model via Markov Chain

The classical spatial models of game dynamics on lattices (73; 83) usually focus on modeling the update rule involving game dynamics with spatial structures such that

$$P \{ Y(s_j, t) \rightarrow Y(s_i, t + 1) \} = f(\pi_j^t, \pi_{N_i}^t) \quad (5.10)$$

where  $f$  is some probability mass function with parameters related to the net/effective payoff  $\pi_j^t$  at location  $s_j$  and  $\pi_{N_i}^t$  in the neighborhood of location  $s_i$  at time  $t$ , respectively.

To incorporate the influence of intensity of selection in the game, Fermi update rule is widely adapted that is defined based on the Fermi distribution with jumping parameter to be a function of the difference between the local payoffs (83) such that

$$P \{ Y(s_j, t) \rightarrow Y(s_i, t + 1) \} = \frac{1}{1 + \exp \left[ -\beta(\pi_j^t - \pi_i^t) \right]}. \quad (5.11)$$

Motivated by (5.10) and (5.11), for the time points  $t \in \{1, 2, \dots, T\}$ , we model the temporal evolution of  $\{\mathbf{Y}_t\}$  by a  $q = |\Omega \setminus \mathcal{I}_t|$ -dimensional vector Markov chain with the transition probability

$$\begin{aligned} & P \left\{ \mathbf{Y}_t(\mathbf{s} \in \Omega \setminus \mathcal{I}_t) \mid \mathbf{Y}_{t-1}(\Omega) \right\} \\ &= \left( \exp [\beta(\boldsymbol{\pi}_D^{t-1} - \boldsymbol{\pi}_C^{t-1}) \circ \mathbf{y}_t(\mathbf{s} \in \Omega \setminus \mathcal{I}_t)] \right) \circ \left( \exp [\beta(\boldsymbol{\pi}_D^{t-1} - \boldsymbol{\pi}_C^{t-1})] + J_{q \times q} \right)^{-1} \end{aligned} \quad (5.12)$$

where  $n \times n$  matrices  $\boldsymbol{\pi}_C^{t-1}$  and  $\boldsymbol{\pi}_D^{t-1}$  are the local average payoff for the cooperators and defectors at the time point  $t - 1$ , respectively;  $\mathbf{Y}_t(\mathbf{s} \in \Omega \setminus \mathcal{I}_t)$  denotes the random variables at active locations at time  $t$  and  $Y_{t-1}(\Omega)$  denotes the process at time point  $t - 1$  on the whole observation domain;  $J_{q \times q}$  is the  $q \times q$  matrix with all component equal to 1; and  $\circ$  is the standard Hadmard product.  $\beta$  is the intensity of selection over time evolution, and can also be interpreted as the conditional variance of the Markov chain along time.

In particular, for each  $s_i \in \Omega \setminus \mathcal{I}_t$

$$P\left\{Y(s_i, t) \mid \mathbf{Y}(N_i, t-1) \cup Y(s_i, t-1)\right\} = \frac{\exp \left[\beta(\boldsymbol{\pi}_D^{i, t-1} - \boldsymbol{\pi}_C^{i, t-1}) y(s_i, t)\right]}{\exp \left[\beta(\boldsymbol{\pi}_D^{i, t-1} - \boldsymbol{\pi}_C^{i, t-1})\right] + 1} \quad (5.13)$$

where  $\boldsymbol{\pi}_C^{i, t-1}$  and  $\boldsymbol{\pi}_D^{i, t-1}$  are the local cooperators' and defector's benefits in the neighborhood of  $s_i$  at time  $t - 1$  as defined in (5.5) and (5.6).

Alternatively, we can rewrite (5.13) into natural exponential form (70; 77) that its Radon-Nikodym derivative is

$$\begin{aligned} & f\left[y(s_i, t) \mid \mathbf{y}(N_i, t-1), y(s_i, t-1)\right] \\ &= \exp \left[A_i(\mathbf{y}(N_i, t-1), y(s_i, t-1)) y(s_i, t) - B(\mathbf{y}(N_i, t-1), y(s_i, t-1))\right] \end{aligned} \quad (5.14)$$

where

$$\begin{aligned} A_i(\mathbf{y}(N_i, t-1), y(s_i, t-1)) &= \beta(\boldsymbol{\pi}_D^{i, t-1} - \boldsymbol{\pi}_C^{i, t-1}) \\ &= \frac{\beta}{|N_i \cup \{s_i\}|} \left( b \sum_{s_j \in N_i \cup \{s_i\}} y(s_j, t-1) - a \sum_{s_j \in N_i \cup \{s_i\}} (1 - y(s_j, t-1)) \right) \\ &= -a\beta + \frac{\beta(a+b)}{|N_i \cup \{s_i\}|} \sum_{s_j \in N_i \cup \{s_i\}} y(s_j, t-1) \end{aligned} \quad (5.15)$$

and

$$B(\mathbf{y}(N_i, t-1), y(s_i, t-1)) = \log \left[1 + \exp \left[A_i(\mathbf{y}(N_i, t-1), y(s_i, t-1))\right]\right]. \quad (5.16)$$

The conditional expectation of  $Y(s_i, t)$  on  $Y(N_i, t-1) \cup Y(s_i, t-1)$  is therefore

$$\begin{aligned} E\left\{Y(s_i, t) \mid \mathbf{Y}(N_i, t-1), y(s_i, t-1)\right\} &= \frac{\exp[A_i(\mathbf{y}(N_i, t-1), y(s_i, t-1))]}{\exp[A_i(\mathbf{y}(N_i, t-1), y(s_i, t-1))] + 1} \\ &= \frac{\exp \left[\beta(\boldsymbol{\pi}_D^{i, t-1} - \boldsymbol{\pi}_C^{i, t-1})\right]}{\exp \left[\beta(\boldsymbol{\pi}_D^{i, t-1} - \boldsymbol{\pi}_C^{i, t-1})\right] + 1} \end{aligned} \quad (5.17)$$

## 5.4 Development of Spatial Dependence Model

We adapt the conditional binary Markov random fields (66; 68; 72; 76) on the compliment of set of inactive locations to model the intrinsic spatial dependence.

Given fixed time point  $t$  and location  $s_i = (u, v) \in \Omega$ , a standard binary Markov random field is formulated by specifying for each  $i = 1, 2, \dots, N$ , a conditional probability mass function in the natural exponential form that

$$f\left[y(s_i, t) \mid \mathbf{y}(N_i, t)\right] = \exp[A_i(\mathbf{y}(N_i, t))y(s_i, t) - B_i(\mathbf{y}(N_i))] \quad (5.18)$$

where

$$A_i(\mathbf{y}(N_i, t)) = \log\left(\frac{\kappa_i}{1 - \kappa_i}\right) + \eta \sum_{s_j \in N_i} \{y(s_j, t) - \kappa_j\} \quad (5.19)$$

and

$$B_i(\mathbf{y}(N_i)) = \log[1 + \exp[A_i(\mathbf{y}(N_i, t))]]. \quad (5.20)$$

As discussed in (77; 79), the natural parameter functions  $A_i(\cdot)$  represent a centered parameterization for binary Markov random field in which, unless the dependence structure/small structure is overwhelmingly large, the parameter  $\kappa_i$  are nearly the marginal expectation of  $Y(s_i, t)$ .

As well know, either the positivity condition of Besag (66) or the Markov random field support condition of Kaiser and Cressie (75) that guarantee the standard model (5.18)-(5.20) to satisfy the condition of existence of a joint distribution that the possesses the specified conditional (66; 72; 75). As discussed in the section 5.1 for the spatial evolutionary game dynamics, both Besag condition and Kaiser-Cressie condition are failed due to the existence of transient forbidden states. We therefore model the spatial dependence via temporal sequences of conditional binary Markov random field based on (5.14)-(5.16) and (5.18)-(5.20) to overcome this difficulty.

Consider a fixed time point  $t$  and  $s_i = (u, v) \in \Omega \setminus \mathcal{I}_t$ , attached to which is its neighborhood  $N_i$ . A binary Markov random field is formulated for the  $|\Omega \setminus \mathcal{I}_t|$  active locations using the conditional distribution (5.18) with natural function  $A_i(\cdot)$  in (5.19) replaced by

$$\begin{aligned} & A_{i,t}(\mathbf{y}(N_i, t), \mathbf{y}(N_i, t-1), y(s_i, t-1)) \\ &= \log\left(\frac{\kappa_{i,t}}{1 - \kappa_{i,t}}\right) + \eta \sum_{s_j \in N_i} \{y(s_j, t) - \kappa_{j,t}\} + \frac{\beta(a+b)}{|N_i \cup \{s_i\}|} \sum_{s_j \in N_i \cup \{s_i\}} y(s_j, t-1) - a\beta \end{aligned} \quad (5.21)$$

where  $|\cdot|$  denotes the cardinality of the set as usual, and  $\eta$  models the strength of spatial dependence while the geometry of  $N_i$  models the structure of spatial dependence.

Conditional on  $\mathbf{Y}_{t-1} = (y(s_1, t-1), y(s_2, t-1), \dots, y(s_N, t-1))$ , the above model specifies a joint distribution for  $\{Y(s_i, t) : s_i \in \Omega \setminus \mathcal{I}_t\}$  as a standard binary Markov random field defined on the lattice specified by the compliment of  $\mathcal{I}_t$  and with neighborhood sets given by  $N_i$ , while unconditionally  $\mathcal{I}_t$  and  $\Omega \setminus \mathcal{I}_t$  are random sets. Hence, for each time point  $t$ , only the locations within the compliment of  $\mathcal{I}_t$  will be updated conditional on the neighborhood information from previous time point  $t-1$  and the neighborhood information at the contemporary time point  $t$ .

In a summary, the complete temporal spatial Markov random field model for evolutionary game dynamics of  $2 \times 2$  games is

$$P\{Y(s_i, 0) | Y(N_i, 0)\} = \frac{\exp[y(s_i, 0)A_i(y(N_i, 0))]}{1 + \exp[y(s_i, 0)A_i(y(N_i, 0))]} \quad (5.22)$$

where

$$A_i(y(N_i, 0)) = \log\left(\frac{\kappa_0}{1 - \kappa_0}\right) + \eta_0 \sum_{s_j \in N_i} (y(s_j, 0) - \kappa_0) \quad (5.23)$$

and with respect to counting measure  $\nu$ , in the natural exponential form

$$\begin{aligned} & P\{Y(s_i, t) | \mathbf{Y}(N_i, t), \mathbf{Y}(N_i, t-1), Y(s_i, t-1)\} \\ &= f\left[y(s_i, t) | \mathbf{y}(N_i, t), \mathbf{y}(N_i, t-1), y(s_i, t-1)\right] \end{aligned} \quad (5.24)$$

$$= \exp[A_i(\mathbf{y}(N_i, t), \mathbf{y}(N_i, t-1), y(s_i, t-1))y(s_i, t) - B_i(\mathbf{y}(N_i), \mathbf{y}(N_i, t-1), y(s_i, t-1))] \quad (5.25)$$

where  $A_{i,t}(\mathbf{y}(N_i, t), \mathbf{y}(N_i, t-1), y(s_i, t-1))$  is defined in (5.21) and

$$B_{i,t}(\mathbf{y}(N_i, t), \mathbf{y}(N_i, t-1), y(s_i, t-1)) = \log[1 + \exp[A_i(\mathbf{y}(N_i, t), \mathbf{y}(N_i, t-1), y(s_i, t-1))]].$$

We assume, for simplicity,  $\kappa_0 = \kappa_i = \kappa$  for each  $i$  and  $\eta_0 = \eta$ .

## 5.5 Existence of Joint Distribution of Spatial-temporal Model

As discussed in (75), specified conditional distribution does not necessarily guarantee the existence of joint distribution. In this section, we will show the existence of the joint probability distribution on the spatial-temporal process with the specified conditional structure proposed above by the method proposed in (79). We have the following result.

**Theorem 5.5.1.** Denote the spatial location by  $\{s_i : i = 1, 2, \dots, N = n^2\}$ , the time points by  $t = 0, 1, 2, \dots, T$ , and denote  $Y(s_i, t) : i = 1, 2, \dots, N; t = 0, 1, \dots, T$  the binary response variables. Let the set of inactive locations at time  $t = 1, 2, \dots, T$  be defined by  $\mathcal{I}_t$  in (5.2). Let the joint distribution of active locations where  $s_i \in \Omega \setminus \mathcal{I}_t$ , conditional on  $N_i$ , be denoted as

$$f\left[\{y(s_i, t) : (s_i, t) \in \Omega \setminus \mathcal{I}_t\}\right] \quad (5.26)$$

where this joint distribution has conditional as in (5.18) with natural functions  $A_{i,t}(\cdot)$  specified by (5.21). Then the probability of a sequence of binary random fields over the time point  $t = 0, 1, 2, \dots, T$  is given by

$$P\left\{\{Y(s_i, t) = y(s_i, t) : i = 1, 2, \dots, N; t = 0, 1, \dots, T\}\right\} = \prod_{t=0}^T f\left[\{y(s_i, t) : (s_i, t) \in \Omega \setminus \mathcal{I}_t\}\right] \quad (5.27)$$

and these probabilities define a distribution for possible sequences.

*Proof.* Let  $\Omega_t$  denote the collection of values after observation of the field at time point  $t = 0, 1, 2, \dots, T$ . Note that there is a duality between the realized values of the field  $\{y(s_i, t) : i = 1, 2, \dots, n\}$  at time  $t$  and  $\Omega_t$ , by which one implies a unique realization of the other and both identify the inactive locations for the next time point and therefore also set of the active locations  $\Omega \setminus \mathcal{I}_{t+1}$ . A sequence of binary fields may therefore be represented as the sequence of collections  $\Omega_0, \Omega_1, \dots, \Omega_T$  such that

$$P\left\{\{Y(s_i, t) = y(s_i, t), i = 1, 2, \dots, N; t = 0, 1, \dots, T\}\right\} = P\left\{\Omega_0, \Omega_1, \dots, \Omega_T\right\}. \quad (5.28)$$

The Markov random field conditional on history, as defined in section 5.4, satisfies the Besag positivity condition (66; 72) so that  $f\left[\{y(s_i, t) : (s_i, t) \in \Omega \setminus \mathcal{I}_t\}\right]$  is well-defined for each  $t$ . Therefore, the probability of  $\Omega_t$  for  $t = 1, \dots, T$  is identified by the duality that

$$P\left\{\Omega_t | \Omega_{t-1}\right\} = f\left[\{y(s_i, t) : (s_i, t) \in \Omega \setminus \mathcal{I}_t\}\right] \quad (5.29)$$

where  $f\left[\{y(s_i, t) : (s_i, t) \in \Omega \setminus \mathcal{I}_t\}\right]$  is the joint distribution of the conditional binary Markov random field defined above.

Notice that the random fields for  $t \geq 1$  depends on the initial configuration  $\Omega_0$  that is well-defined by its form (5.7)-(5.8) and Besag positivity condition (66; 72). The neighborhood collections  $\Omega_t$  follow

a Markov property as specified in section 5.3 that

$$\begin{aligned} P\{\Omega_0, \Omega_1, \dots, \Omega_T\} &= P(\Omega_0)P(\Omega_1|\Omega_0)\cdots P(\Omega_T|\Omega_{T-1}) \\ &= \prod_{t=0}^T f\left[\{y(s_i, t) : (s_i, t) \in \Omega \setminus \mathcal{I}_t\}\right]. \end{aligned} \quad (5.30)$$

Let  $\mathcal{C}_{\Omega_t|\Omega_{t-1}}$  denote the set of possible configurations for  $\Omega_t$  given  $\Omega_{t-1}$ . The conditional probabilities (5.29) are defined by the well-posed joint distributions and each element of  $\mathcal{C}_{\Omega_t|\Omega_{t-1}}$  uniquely corresponds to a value in the support of these distribution which is defined by

$$\mathcal{S}_t = \{\text{permutations of } \{0, 1\}^{|\Omega \setminus \mathcal{I}_t|}\}.$$

Notice, for each  $t = 1, 2, \dots, T$ , with respect to counting measure,

$$\sum_{\mathcal{C}_{\Omega_t|\Omega_{t-1}}} P\{\Omega_t | \Omega_{t-1}\} = \int_{\mathcal{S}_t} f\left[\{y(s_i, t) : (s_i, t) \in \Omega \setminus \mathcal{I}_t\}\right] d\nu = 1. \quad (5.31)$$

Denote the set of all possible configurations of the sequence of binary Markov random fields over the time  $t = 0, 1, 2, \dots, T$  by  $\mathcal{C}_T$ , which is a joint union of Cartesian products of conditional possible configurations that

$$\mathcal{C}_T = \bigcup \{\mathcal{C}_{\Omega_0} \times \mathcal{C}_{\Omega_1|\Omega_0} \times \mathcal{C}_{\Omega_2|\Omega_1} \times \cdots \times \mathcal{C}_{\Omega_T|\Omega_{T-1}}\}.$$

Hence, by (5.30) and (5.31), the total probabilities for sequence of random field configurations over all possible configurations in  $\mathcal{C}_T$  is

$$\begin{aligned} &\sum_{\mathcal{C}_T} P\{\Omega_0, \Omega_1, \dots, \Omega_T\} \\ &= \sum_{\mathcal{C}_{\Omega_0} \times \mathcal{C}_{\Omega_1|\Omega_0} \times \mathcal{C}_{\Omega_2|\Omega_1} \times \cdots \times \mathcal{C}_{\Omega_T|\Omega_{T-1}}} P(\Omega_0)P(\Omega_1|\Omega_0)\cdots P(\Omega_T|\Omega_{T-1}) \\ &= \sum_{\mathcal{C}_{\Omega_0}} P(\Omega_0) \sum_{\mathcal{C}_{\Omega_1|\Omega_0}} P(\Omega_1|\Omega_0) \cdots \sum_{\mathcal{C}_{\Omega_T|\Omega_{T-1}}} P(\Omega_T|\Omega_{T-1}) \\ &= \sum_{\mathcal{C}_{\Omega_0}} P(\Omega_0) \\ &= 1 \end{aligned} \quad (5.32)$$

by the initial configuration's well-posedness. The assertion is therefore derived.  $\square$

## 5.6 Future Work

We introduce the main idea on development of Markov Random Field Model, which incorporates replicator dynamics and mainly deals with observations on a lattice using appropriate random variable distribution, to estimate the parameters of the payoff matrix for linear game model. In the near future, a simulation study will be explored and study of different games with respect to planes consist of paired parameters will be implemented. And the parameter estimation procedure will be a process of maximization of the Besag likelihood function over the parameter space for the inverse spatial game, and then the inference of the parameters will be based a MCMC method. Furthermore, model assessment approaches will be followed to justify the performance of the proposed Markov Random Field model. Finally we will investigate the spatial effects of interaction region or neighborhood size on the estimated parameters for given data and analyze the nature of the game that we are interested in. It is expected that the Markov Random Field model will be useful in applying to a lot of spatial-temporal studies consist of two species or subpopulations in biological or social systems.

## CHAPTER 6. OPTIMALITY AND STABILITY OF SYMMETRIC EVOLUTIONARY GAMES WITH APPLICATIONS IN GENETIC SELECTION

A paper published in *Journal of Mathematical Biosciences and Engineering*

Yuanyuan Huang, Yiping Hao, Min Wang, Wen Zhou, Zhijun Wu

### **Abstract**

Symmetric evolutionary games, i.e., evolutionary games with symmetric fitness matrices, have important applications in population genetics, where they can be used to model for example the selection and evolution of the genotypes of a given population. In this paper, we review the theory for obtaining optimal and stable strategies for symmetric evolutionary games, and provide some new proofs and computational methods. In particular, we review the relationship between the symmetric evolutionary game and the generalized knapsack problem, and discuss the first and second order necessary and sufficient conditions that can be derived from this relationship for testing the optimality and stability of the strategies. Some of the conditions are given in different forms from those in previous work and can be verified more efficiently. We also derive more efficient computational methods for the evaluation of the conditions than conventional approaches. We demonstrate how these conditions can be applied to justifying the strategies and their stabilities for a special class of genetic selection games including some in the study of genetic disorders.

### **6.1 Introduction**

We consider an  $n$ -strategy evolutionary game defined by a symmetric fitness matrix  $A \in R^{n \times n}$ . Let  $S = \{x \in R^n : x \geq 0, \sum_i x_i = 1\}$  be the set of all mixed strategies. The problem is to find an optimal

strategy  $x^* \in S$  such that

$$x^{*T} Ax^* \geq x^T Ax^* \text{ for all } x \in S. \quad (6.1)$$

We call this problem a symmetric evolutionary game or SEgame for short. The problem has important applications in population genetics, where it can be used to model and study the evolution of genotypes in a given population when their corresponding phenotypes are under selection pressures.

The modeling of genetic selection has a long history (89). It may be traced back to the earliest mathematical work in population genetics in early last century including the Hardy-Weinberg's Law by G. H. Hardy and W. Weinberg in 1908 (91; 103) and the Fundamental Theorem of Natural Selection by R. A. Fisher in 1930 (90). The work has especially been revived in 1970s when J. Maynard Smith introduced the game theory to biology and developed the evolutionary game theory for the study of evolution of population of competing species (93). In this theory, a genetic selection problem can in particular be modeled as a SEgame (92).

The SEgame has a close relationship with the generalized knapsack problem or GKproblem for short, which is to find an optimal solution  $x^* \in R^n$  for the following maximization problem:

$$\max_{x \in R^n} x^T Ax/2 \quad (6.2)$$

$$\text{subject to } \sum_i x_i = 1, x \geq 0.$$

The GKproblem has been studied extensively, with applications in solving maximum clique problems (94), in convex quadratic programming (98), and especially in game theoretic modeling (85).

In this paper, we review the theory for obtaining optimal and stable strategies for symmetric evolutionary games, and provide some new proofs and computational methods. In particular, we review the relationship between the symmetric evolutionary game and the generalized knapsack problem, and discuss the first and second order necessary and sufficient conditions that can be derived from this relationship for testing the optimality and stability of the strategies. Some of the conditions are given in different forms from those in previous work and can be verified more efficiently. We also derive more efficient computational methods for the evaluation of the conditions than conventional approaches. We demonstrate how these conditions can be applied to justifying the strategies and their stabilities for a special class of genetic selection games including some in the study of genetic disorders.

### 6.1.1 Further Mathematical Background

A two-player game is said to be symmetric if the players share the same fitness matrix and the same set of strategies. Let  $A \in R^{n \times n}$  be the fitness matrix and  $S = \{x \in R^n : x \geq 0, \sum_i x_i = 1\}$  the set of all mixed strategies. Let  $x \in S$  be the strategy played by player I and  $y \in S$  by player II. Then, the fitness for player I can be defined by a function  $\pi(x, y) = x^T A y$  and for player II by  $\pi(y, x) = y^T A x$ . A pair of strategies  $(x^*, y^*)$  is said to be optimal if  $x^{*T} A y^* \geq x^T A y^*$  for all  $x \in S$  and  $y^{*T} A x^* \geq y^T A x^*$  for all  $y \in S$ , where  $x^*$  and  $y^*$  are said to be the best response to each other (see Fig. 6.1).

|       |           |           |
|-------|-----------|-----------|
|       | $y_1$     | $y_2$     |
| $x_1$ | $A_{1,1}$ | $A_{1,2}$ |
| $x_2$ | $A_{2,1}$ | $A_{2,2}$ |

Figure 6.1 Two-Player Game: A two-player, two-strategy symmetric game is demonstrated. The strategies for player I are given in vector  $x = (x_1, x_2)^T$ , and for player II in  $y = (y_1, y_2)^T$ ,  $x, y \in S = \{x \in R^2 : \sum_i x_i = 1, x_i \geq 0, i = 1, 2\}$ . The fitness  $A_{i,j}$  of strategy pair  $(x_i, y_j)$  is given in the  $(i, j)$ -entry of a  $2 \times 2$  fitness matrix  $A$ . A strategy pair  $(x^*, y^*)$  is said to be optimal if  $x^{*T} A y^* \geq x^T A y^*$  for all  $x \in S$  and  $y^{*T} A x^* \geq y^T A x^*$  for all  $y \in S$ , when the game is said to reach the Nash equilibrium.

A special class of symmetric games is to find a strategy  $x^* \in S$  which is the best response to itself, i.e., player I and II play the same strategy  $x^*$  and  $x^{*T} A x^* \geq x^T A x^*$  for all  $x \in S$ . This class of games is often used to model the evolution of a population of competing species, with player I being a particular individual and player II being a typical individual in the population. A strategy  $x$  for player I means the species type the particular individual prefers to be. It could be a pure species type, i.e.,  $x = e_i$  for some  $i$  or a mixed one with  $x_i \neq 1$  for any  $i$ , where  $e_i$  is the  $i$ th unit vector. Note that by a mixed species type  $x$  we mean the frequency of the individual to play species  $i$  is  $x_i$ . On the other hand, a strategy  $y$  for player II means the typical species type of an individual in the population, which depends on the species composition of the population. More specifically, if the portion for species  $i$  in

the population is  $y_i$ , then the chance for a typical individual to be species  $i$  is also  $y_i$ . Therefore,  $y$  is also a population profile, and  $x^T A y$  is basically the fitness for species  $x$  in population  $y$ . Such a game is called a population game, or an evolutionary game, or a game against the field (99; 102). The goal of the game is to find an optimal strategy  $x^* \in S$  so that in population  $x^*$ , an individual cannot find a better strategy than  $x^*$ , i.e.,  $x^{*T} A x^* \geq x^T A x^*$  for all  $x \in S$ , which is when the population has reached the so-called Nash equilibrium. Biologically, this is when the population has reached a state so that the optimal strategy for an individual is a species type consistent with the typical species type of the population. If the fitness matrix of a symmetric game itself is symmetric, the game is called a doubly symmetric game (102). An evolutionary game with a symmetric fitness matrix is a doubly symmetric game, which is what we call a symmetric evolutionary game, i.e., a SEgame as given in (6.1).

### 6.1.2 Further Biological Background

SEgames can be used to model genetic selection and in particular, allele selection. An allele is one of several possible forms of a gene. Most of multi-cellular organisms are diploid, i.e., their chromosomes form homologous pairs. Each pair of chromosomes has a pair of alleles at each genetic locus. Thus,  $n$  different alleles may form  $n^2$  different allele pairs, as two alleles in each pair may not be the same. Different allele pairs are considered to be different genotypes, which may result in different phenotypes or in other words, different genetic traits (see Fig. 6.2).

The fitness of all different allele pairs or in other words, all different genotypes at a given genetic locus can then be given in a matrix with the rows corresponding to the choices for the first allele and the columns to the choices for the second allele in the allele pair. Again,  $n$  different alleles will give  $n$  different choices for both the first and second alleles in the allele pair, and hence an  $n \times n$  fitness matrix. With such a fitness matrix, a genetic selection game can then be defined with the choices of the first and second alleles in the allele pair at a given genetic locus as the strategies for player I and II. Here, player I can be considered as an individual with a specific choice of allele at the given locus. The choice could be one of the possible alleles or a combination of them with each selected with some chance. The former corresponds to a pure strategy, while the latter to a mixed one. In any case, if there are  $n$  different alleles, the strategy for player I can be represented by a vector  $x \in R^n$ ,  $x \geq 0$ ,  $\sum_i x_i = 1$ . On the other hand, player II can be considered as a typical individual in the given population. This

|     |   |   |     |
|-----|---|---|-----|
|     | $A$   |   | $a$ |
| $A$ |  |  |     |
| $a$ |  |  |     |

Figure 6.2 Genetic Selection: In diploid species, there are always two alleles at each genetic locus. Each pair of alleles determines a certain genotype, which in turn determines a certain phenotype. For example, in Wendel's classical experiment, the color of the flowers depends on the pairing of the alleles at a specific genetic locus, one for pink color and dominant, and another for white and recessive. Let the dominant allele be denoted by  $A$  and the recessive one by  $a$ . There can be four possible allele pairs,  $AA$ ,  $Aa$ ,  $aA$ , and  $aa$ . Since  $A$  is dominant,  $AA$ ,  $Aa$ , and  $aA$  will produce pink flowers, while  $aa$  will produce white ones. These genotypic and phenotypic outcomes can be summarized in a  $2 \times 2$  allele-pairing matrix as arranged in the figure.

individual could have only one of possible alleles at the given locus or a combination of them with each selected with some chance. Similar to player I, if there are  $n$  different alleles, the strategy for player II can be represented by a vector  $y \in R^n$ ,  $y \geq 0$ ,  $\sum_i y_i = 1$ . This strategy  $y$  really is the same as the composition of alleles at the given locus in the whole population. Therefore, it is also the allele profile of the population for this particular genetic locus. Let the fitness matrix be given by  $A \in R^{n \times n}$ . Let  $S = \{x \in R^n : x \geq 0, \sum_i x_i = 1\}$ . The average fitness of an allele choice  $x \in S$  in an allele population  $y \in S$  will be  $x^T A y$ . We then want to find an optimal choice of  $x^* \in S$  such that  $x^{*T} A x^* \geq x^T A x^*$  for all  $x \in S$ , i.e., in allele population  $x^*$ , any individual with allele choice  $x$  other than  $x^*$  will not have a better average fitness than allele choice  $x^*$  (92). Note that the fitness for allele pair  $(i, j)$  usually is the same as that for  $(j, i)$ . Therefore, the fitness matrix for genetic selection is typically symmetric, and the corresponding game is then a SGame.

## 6.2 GKproblems vs. SGames

For an evolutionary game, it is well known that a mixed strategy  $x^* \in S$  is optimal for the game if and only if the fitness  $x^{*T} A x^* = (Ax^*)_i$  for all  $i$  such that  $x_i^* > 0$  and  $x^{*T} A x^* \geq (Ax^*)_i$  for all  $i$

such that  $x_i^* = 0$  (99; 102). These conditions also apply to any symmetric evolutionary game, i.e., any SEgame in (6.1), and can be stated formally as in the following theorem.

**Theorem 6.2.1.** *Let  $A \in R^{n \times n}$  be a symmetric fitness matrix and  $S = \{x \in R^n : x \geq 0, \sum_i x_i = 1\}$  the set of all mixed strategies. Then, a strategy  $x^* \in S$  is an optimal strategy for the SEgame in (6.1) if and only if there is a scalar  $\lambda^*$  such that*

$$x_i^* \geq 0, \lambda^* - (Ax^*)_i \geq 0 \quad (6.3)$$

$$x_i^*(\lambda^* - (Ax^*)_i) = 0, i = 1, \dots, n. \quad (6.4)$$

The proof of the above theorem can be found in many text books such as in (99; 102). Since it is helpful for the understanding of the nature of the optimal strategies of the SEgame, we also provide one here for the self-containedness of the paper:

*Proof.* If  $x^* \in S$  satisfies the conditions in (6.3) and (6.4), by adding all equations in (6.4), we then obtain  $\lambda^* = x^{*T} Ax^*$ . Let  $x \in S$  be an arbitrary strategy. Multiply the second inequality in (6.3) by  $x_i$ . Then, by adding all second inequalities in (6.3), we obtain  $\lambda^* - x^T Ax^* \geq 0$ , i.e.,  $x^{*T} Ax^* \geq x^T Ax^*$ , since  $\lambda^* = x^{*T} Ax^*$ . Therefore,  $x^*$  is an optimal strategy for the SEgame in (6.1).

If  $x^* \in S$  is an optimal strategy for the SEgame in (6.1), then  $x^{*T} Ax^* \geq x^T Ax^*$  for any  $x \in S$  and therefore,  $x^{*T} Ax^* \geq e_i^T Ax^* = (Ax^*)_i$  for all  $i$ . Let  $\lambda^* = x^{*T} Ax^*$ . Then,  $\lambda^* - (Ax^*)_i \geq 0$  for all  $i$ . Assume that  $x_i^*(\lambda^* - (Ax^*)_i) > 0$  for some  $i$ . By adding all the left-hand sides of the equations in (6.4), we then obtain  $\lambda^* > x^{*T} Ax^*$ , which contradicts to the fact that  $\lambda^* = x^{*T} Ax^*$ . Therefore,  $x_i^*(\lambda^* - (Ax^*)_i) = 0$  for all  $i$ .  $\square$

As we have mentioned in Section 6.1, the symmetric evolutionary game, i.e., the SEgame in (6.1) is closely related to the generalized knapsack problem, i.e., the GKproblem in (6.2). A knapsack problem is originally referred to as a problem for selecting a set of objects of different sizes and values into a given sack of fixed size to maximize the total value of objects in the sack. The problem can be formulated as a linear program, with a linear objective function  $\sum_i a_i x_i$  for the total value of the sack, where  $x_i$  and  $a_i$  are the size and unit value of object  $i$ , respectively and with a linear constraint  $\sum_i x_i \leq s$ ,  $x_i \geq 0$ ,  $i = 1, \dots, n$  on the total size of the objects that can be put into the sack, where  $n$  is the number of objects and  $s$  the size of the sack. The GKproblem in (6.2) can therefore be considered as

a knapsack problem of  $n$  “objects” with the objective function generalized to a symmetric quadratic form  $x^T Ax/2$  and with the “sack” restricted in a simplex  $S = \{x \in R^n : x \geq 0, \sum_i x_i = 1\}$ . If we interpret the “objects” to be the species fractions in a given population and the matrix  $A$  to be the fitness matrix of the species, the objective function for the GKproblem in (6.2) is exactly half of the average fitness of the population of the SGame in (6.1). Therefore, the goal of the GKproblem in (6.2) is basically to maximize the average fitness of the population of the SGame in (6.1).

Based on general optimization theory, an optimal solution to the GKproblem in (6.2) must satisfy certain conditions. We first consider a general constrained optimization problem

$$\min_{x \in R^n} f(x) \quad (6.5)$$

$$\text{subject to } c_i(x) = 0, i \in E$$

$$c_i(x) \geq 0, i \in I$$

where  $f(x)$  is the objective function,  $c_i(x)$  the constraint functions,  $E$  the set of indices for equality constraints, and  $I$  the set of indices for inequality constraints. Assume that  $f(x)$  and  $c_i(x)$  are all continuously differentiable. Let  $x$  be a feasible solution for the problem, i.e.,  $c_i(x) = 0, i \in E$  and  $c_i(x) \geq 0, i \in I$ . Let  $E_0(x)$  be the set of indices for the constraints active at  $x$ , i.e.,  $E_0(x) = E \cup \{i \in I : c_i(x) = 0\}$  and  $C_0(x)$  be the Jacobian of the constraints active at  $x$ , i.e.,  $C_0(x) = \{\nabla c_i(x) : i \in E_0(x)\}^T$ . We then have a set of first-order necessary conditions for an optimal solution to the general constrained optimization problem in (6.5) as can be stated in the following theorem. Here, we say that  $x^* \in R^n$  is an optimal solution for the general constrained optimization problem in (6.5), if  $x^*$  is feasible, i.e.,  $x^*$  satisfies all the constraints, and if  $f(x^*) \leq f(x)$  for all  $x$  feasible in a small neighborhood  $U$  of  $x^*$ .

**Theorem 6.2.2** ((97)). *Let  $x^* \in R^n$  be an optimal solution to the general constrained optimization problem in (6.5). Assume that the gradients of the constraints active at  $x^*$ , i.e., the vectors in  $C_0(x^*)$ , are linearly independent. Then, there must be a set of Lagrange multipliers  $\lambda^* \in R^{|E|}$  and  $\mu^* \in R^{|I|}$*

such that

$$\begin{aligned}
\nabla_x L(x^*, \lambda^*, \mu^*) &= 0, \\
c_i(x^*) &= 0, \quad i \in E, \\
c_i(x^*) &\geq 0, \quad i \in I, \\
\mu_i^* &\geq 0, \quad i \in I, \\
\mu_i^* c_i(x^*) &= 0, \quad i \in I,
\end{aligned} \tag{6.6}$$

where  $L(x, \lambda, \mu)$  is called the Lagrangian function of the problem in (6.5),

$$L(x, \lambda, \mu) = f(x) - \sum_{i \in E} \lambda_i c_i(x) - \sum_{i \in I} \mu_i c_i(x).$$

The conditions in (6.6) are called the KKT conditions of the general constrained optimization problem in (6.5) named after W. Karush, H. Kuhn, and A. Tucker, who first discovered and proved the conditions. As stated in Theorem 6.2.2, an optimal solution  $x^*$  of the general constrained optimization problem in (6.5) must satisfy the KKT conditions, but a feasible solution  $x^*$  that satisfies the KKT conditions, called a KKT point, may not always be an optimal solution.

We now apply Theorem 6.2.2 to the GKproblem in (6.2). By changing the maximization problem to a standard minimization problem, we then have the objective function for this problem  $f(x) = -x^T A x / 2$ . If we name the nonnegative constraints  $c_i(x) = x_i \geq 0$ ,  $i = 1, \dots, n$  to be the first to the  $n$ th constraints and the equality constraint  $c_{n+1}(x) = 1 - \sum_i x_i = 0$  to be the  $n+1$ th constraint, we then have  $I = \{1, \dots, n\}$  and  $E = \{n+1\}$ . Let  $x$  be a feasible solution for the problem. Let  $E_0(x)$  be the set of indices for the constraints active at  $x$ , i.e.,  $E_0(x) = \{i \in I : c_i(x) = 0\} \cup E$  and  $C_0(x)$  be the Jacobian of the constraints active at  $x$ , i.e.,  $C_0(x) = \{\nabla c_i(x) : i \in E_0(x)\}^T$ . Then,  $E_0(x) = \{i \in I : x_i = 0\} \cup \{n+1\}$  and  $C_0(x)^T = \{e_i : i \in I, x_i = 0\} \cup \{-e\}$ , where  $e_i$  is the  $i$ th unit vector and  $e = \sum_i e_i$ . For any  $x \in S$ , there is at least one  $i \in I$  such that  $x_i \neq 0$  since  $x \geq 0$  and  $\sum_i x_i = 1$ . Therefore,  $E_0$  includes the index  $n+1$  and a subset of indices  $\{i \in I\}$ , and  $C_0(x)$  contains the vector  $-e^T$  and a subset of vectors  $\{e_i^T : i \in I\}$ , which are always linearly independent. We then have the following first-order necessary conditions for the GKproblem in (6.2):

**Theorem 6.2.3.** *Let  $A \in R^{n \times n}$  be a symmetric fitness matrix and  $S = \{x \in R^n : x \geq 0, \sum_i x_i = 1\}$  the set of all feasible solutions for the GKproblem in (6.2). If  $x^* \in S$  is an optimal solution for this*

problem, then there must be a scalar  $\lambda^*$  such that

$$x_i^* \geq 0, \lambda^* - (Ax^*)_i \geq 0, \quad (6.7)$$

$$x_i^*(\lambda^* - (Ax^*)_i) = 0, \quad i = 1, \dots, n. \quad (6.8)$$

*Proof.* The Lagrangian function for the GKproblem in (6.2) can be written in the following form:

$$L(x, \lambda, \mu) = -x^T Ax/2 - \lambda(1 - \sum_i x_i) - \mu^T x.$$

where  $x \in R^n$ ,  $\lambda \in R$ ,  $\mu \in R^n$ . Since for this problem the gradients of the active constraints at any  $x \in S$ , i.e., the vectors in  $C_0(x)$ , are linearly independent, by Theorem 6.2.2, if  $x^* \in S$  is an optimal solution to the GKproblem in (6.2), then there must be  $\lambda^* \in R$ ,  $\mu^* \in R^n$  such that

$$\begin{aligned} -Ax^* + \lambda^*e - \mu^* &= 0, \\ \sum_i x_i^* &= 1, \\ x^* &\geq 0, \\ \mu^* &\geq 0, \\ x^{*T} \mu^* &= 0. \end{aligned}$$

By substituting  $\mu^* = \lambda^*e - Ax^*$  in all the formulas, we then have

$$\begin{aligned} x^* &\geq 0, \lambda^*e - Ax^* \geq 0, \\ x^{*T}(\lambda^*e - Ax^*) &= 0, \end{aligned}$$

which are equivalent to the conditions in (6.7) and (6.8).  $\square$

Note that the conditions in (6.3) and (6.4) of Theorem 6.2.1 and in (6.7) and (6.8) of Theorem 6.2.3 are the same. However, it does not imply that the SEgame in (6.1) is equivalent to the GKproblem in (6.2), because the conditions are necessary and sufficient for an optimal strategy for the SEgame in (6.1) but only necessary for an optimal solution for the GKproblem in (6.2). Therefore, an optimal solution for the GKproblem in (6.2) must be an optimal strategy for the SEgame in (6.1), while the converse may not necessarily be true. We state this conclusion as a corollary from Theorem 6.2.1 and 6.2.3 in the following.

**Corollary 6.2.4.** *An optimal solution  $x^* \in S$  for the GKproblem in (6.2) must be an optimal strategy for the SEgame in (6.1), while an optimal strategy  $x^* \in S$  for the SEgame in (6.1) is only a KKT point for the GKproblem in (6.2), which is necessary but not sufficient to be optimal for the GKproblem in (6.2).*

In any case, the above two types of problems are closely related. The properties of the optimal strategies for a SEgame can be investigated by examining the nature of the optimal solutions to the corresponding GKproblem. For example, the existence of the optimal strategy for a general game, which usually requires a more involved theoretical proof (96), now becomes much easier to verify for a SEgame based on the relationship between the SEgame and the GKproblem: There is always an optimal solution for the GKproblem in (6.2), given the fact that the objective function of the problem is a continuous function and the feasible set is a bounded and closed simplex. Based on Corollary 6.2.4, an optimal solution for the GKproblem in (6.2) is an optimal strategy for the SEgame in (6.1). Then, the next corollary follows:

**Corollary 6.2.5.** *There is always an optimal strategy or in other words, a Nash equilibrium for a given SEgame in (6.1).*

The fact that an optimal strategy for the SEgame in (6.1) maximizes the objective function of the GKproblem in (6.2) has been recognized in (99; 102) and discussed in great detail in (85). However, they have focused on the equivalence between the two types of problems when the strategy is evolutionarily stable, weak or strong. Here, we have made a clear distinction between them and shown that the strategies for the SEgame in (6.1) are not necessarily always be optimal solutions of the GKproblem in (6.2). When not, they can be local minimizers or saddle points of the GKproblem in (6.2). Though unstable, they can be interesting to analyze as well, as we will mention again in our concluding remarks in Section 6.8. Besides, we have provided detailed proofs for the necessary and sufficient conditions for both types of problems. Based on these proofs, we have been able to obtain the Corollary 6.2.5 easily for the existence of the equilibrium state of the SEgame in (6.1).

### 6.3 Second-order Optimality Conditions

We now focus on the GKproblem in (6.2) and derive additional second-order necessary and sufficient conditions for its optimal solutions, and extend them to the solutions for the SEgame in (6.1). These conditions have been mentioned in several literature (99; 102) and especially analyzed in great detail in (85). Here we review the conditions, with some given in different forms from those in (85). They are in fact weaker conditions, but easier to verify, which is important for the later development of our computational methods for justifying the solutions and their stabilities for the GKproblems as well as the SEgames. We will comment more on these differences in the end of this section.

Consider again the general constrained optimization problem in (6.5). Let  $x^*$  be an optimal solution to the problem. Let  $E_0(x^*)$  be the set of indices for the constraints active at  $x^*$ , i.e.,  $E_0(x^*) = E \cup \{i \in I : c_i(x^*) = 0\}$  and  $C_0(x^*)$  be the Jacobian of the constraints active at  $x^*$ , i.e.,  $C_0(x^*) = \{\nabla c_i(x^*) : i \in E_0(x^*)\}^T$ . We then have the following second-order necessary conditions for  $x^*$  to be an optimal solution to the problem in (6.5).

**Theorem 6.3.1** ((97)). *Let  $x^* \in R^n$  be an optimal solution to the general constrained optimization problem in (6.5). Assume that  $C_0(x^*)$  has full row rank  $m$ . Let  $Z_0 \in R^{n \times (n-m)}$  be the null space matrix of  $C_0(x^*)$ . Then,*

$$y^T Z_0^T \nabla^2 f(x^*) Z_0 y \geq 0 \text{ for all } y \in R^{n-m}, y \neq 0, \quad (6.9)$$

i.e., the reduced Hessian of  $f(x)$  at  $x^*$ ,  $Z_0^T \nabla^2 f(x^*) Z_0$ , must be positive semi-definite.

Now consider a KKT point  $x^* \in R^n$  for the general constrained optimization problem in (6.5). Let  $E^0(x^*)$  be the set of indices for the constraints strongly active at  $x^*$ , i.e.,  $E^0(x^*) = E \cup \{i \in I : c_i(x^*) = 0 \text{ and } \mu_i^* > 0\}$  and  $C^0(x^*)$  be the Jacobian of the constraints strongly active at  $x^*$ , i.e.,  $C^0(x^*) = \{\nabla c_i(x^*) : i \in E^0(x^*)\}^T$ , where  $\mu_i^*$  are the Lagrangian multipliers for the inequality constraints in the KKT conditions. We then have the following second-order sufficient conditions for  $x^*$  to be a strict optimal solution to the problem in (6.5), i.e.,  $f(x^*) < f(x)$  for all feasible solution  $x \neq x^*$  in some neighborhood  $U$  of  $x^*$ .

**Theorem 6.3.2** ((97)). *Let  $x^* \in R^n$  be a KKT point for the general constrained optimization problem in (6.5). Assume that  $C^0(x^*)$  has full row rank  $m$ . Let  $Z^0 \in R^{n \times (n-m)}$  be the null space matrix of*

$C^0(x^*)$ . If

$$y^T Z^{0T} \nabla^2 f(x^*) Z^0 y > 0 \text{ for all } y \in R^{n-m}, y \neq 0, \quad (6.10)$$

i.e., the reduced Hessian of  $f(x)$  at  $x^*$ ,  $Z^{0T} \nabla^2 f(x^*) Z^0$ , is positive definite, then  $x^*$  must be a strict optimal solution to the problem in (6.5).

We now apply Theorem 6.3.1 and 6.3.2 to the GKproblem in (6.2). By changing the maximization problem to a standard minimization problem, we then have the objective function for the GKproblem in (6.2) to be  $f(x) = -x^T A x / 2$ . If we name the nonnegative constraints  $c_i(x) = x_i \geq 0$ ,  $i = 1, \dots, n$  to be the first to the  $n$ th constraints and the equality constraint  $c_{n+1}(x) = 1 - \sum_i x_i = 0$  to be the  $n+1$ th constraint, we then have  $I = \{1, \dots, n\}$  and  $E = \{n+1\}$ . Let  $x^* \in S$  be a KKT point for the GKproblem in (6.2). Let  $E_0(x^*)$  be the set of indices for the constraints active at  $x^*$ , i.e.,  $E_0(x^*) = \{i \in I : c_i(x^*) = 0\} \cup E$  and  $C_0(x^*)$  be the Jacobian of the constraints active at  $x^*$ , i.e.,  $C_0(x^*) = \{\nabla c_i(x^*) : i \in E_0(x^*)\}^T$ . Then,  $E_0(x^*) = \{i \in I : x_i^* = 0\} \cup \{n+1\}$  and  $C_0(x^*)^T = \{e_i : i \in I, x_i^* = 0\} \cup \{-e\}$ , where  $e_i$  is the  $i$ th unit vector and  $e = \sum_i e_i$ . For any  $x^* \in S$ , there is at least one  $i \in I$  such that  $x_i^* \neq 0$  since  $x^* \geq 0$  and  $\sum_i x_i^* = 1$ . Therefore,  $E_0$  includes the index  $n+1$  and a subset of indices  $\{i \in I\}$ , and  $C_0(x^*)$  contains the vector  $-e^T$  and a subset of vectors  $\{e_i^T : i \in I\}$  as the rows, and is of full row rank. Note also that the Hessian of the objective function  $\nabla^2 f(x^*) = -A$ . We then have the following second-order necessary conditions for  $x^*$  to be an optimal solution to the GKproblem in (6.2).

**Theorem 6.3.3.** Let  $x^* \in S$  be an optimal solution to the GKproblem in (6.2). Let the row rank of  $C_0(x^*)$  be equal to  $m$ , and  $Z_0 \in R^{n \times (n-m)}$  the null space matrix of  $C_0(x^*)$ . Then,

$$y^T Z_0^T A Z_0 y \leq 0 \text{ for all } y \in R^{n-m}, y \neq 0, \quad (6.11)$$

i.e., the reduced Hessian of the objective function of the GKproblem in (6.2) at  $x^*$ ,  $Z_0^T A Z_0$ , must be negative semi-definite.

Now consider a KKT point  $x^* \in S$ . Let  $E^0(x^*)$  be the set of indices for the constraints strongly active at  $x^*$ , i.e.,  $E^0(x^*) = \{i \in I : c_i(x^*) = 0 \text{ and } \mu_i^* > 0\} \cup E$  and  $C^0(x^*)$  be the Jacobian of the constraints strongly active at  $x^*$ , i.e.,  $C^0(x^*) = \{\nabla c_i(x^*) : i \in E^0(x^*)\}^T$ , where  $\mu_i^*$  are the Lagrangian multipliers for the inequality constraints in the KKT conditions for the GKproblem in (6.2),

$\mu_i^* = x^{*T}Ax^* - (Ax^*)_i$  (see in the proof for Theorem 6.2.3). Then,  $E^0(x^*) = \{i \in I : x_i^* = 0 \text{ and } \mu_i^* > 0\} \cup \{n+1\}$  and  $C^0(x^*)^T = \{e_i : i \in I, x_i^* = 0 \text{ and } \mu_i^* > 0\} \cup \{-e\}$ , where  $e_i$  is the  $i$ th unit vector and  $e = \sum_i e_i$ . Again, for any  $x^* \in S$ , there is at least one  $i \in I$  such that  $x_i^* \neq 0$  since  $x^* \geq 0$  and  $\sum_i x_i^* = 1$ . Therefore,  $E^0$  includes the index  $n+1$  and a subset of indices  $\{i \in I\}$ , and  $C^0(x^*)$  contains the vector  $-e^T$  and a subset of vectors  $\{e_i^T : i \in I\}$  as rows, and is of full row rank. Note also that the Hessian of the objective function  $\nabla^2 f(x^*) = -A$ . We then have the following second-order sufficient conditions for  $x^*$  to be a strict optimal solution to the GKproblem in (6.2).

**Theorem 6.3.4.** *Let  $x^* \in S$  be a KKT point for the GKproblem in (6.2). Let the row rank of  $C^0(x^*)$  be equal to  $m$ . Let  $Z^0 \in R^{n \times (n-m)}$  be the null space matrix of  $C^0(x^*)$ . Then  $x^*$  must be a strict optimal solution to the GKproblem in (6.2) if*

$$y^T Z^{0T} A Z^0 y < 0 \text{ for all } y \in R^{n-m}, y \neq 0, \quad (6.12)$$

i.e., the reduced Hessian of the objective function of the GKproblem in (6.2) at  $x^*$ ,  $Z^{0T} A Z^0$ , is negative definite.

Note that the conditions in Theorem 6.3.3 and 6.3.4 are either necessary or sufficient but not both. In fact, since the GKproblem in (6.2) is a quadratic program, it is possible to establish a second-order necessary and sufficient condition for its optimal solution. For this purpose, we go back to the general constrained optimization problem (6.5) again. Let  $x \in R^n$  be any feasible solution for the problem. We define the reduced tangent cone  $T(x)$  at  $x$  to be the set of vectors  $d \in R^n$  such that

$$\nabla c_i(x)^T d = 0, \text{ for all } i \in E, \quad (6.13)$$

$$\nabla c_i(x)^T d = 0, \text{ for all } i \in I \text{ such that } c_i \text{ strongly active at } x, \quad (6.14)$$

$$\nabla c_i(x)^T d \geq 0, \text{ for all } i \in I \text{ such that } c_i \text{ weakly active at } x. \quad (6.15)$$

Then, based on general optimization theory, we know that if the general constrained optimization problem in (6.5) is a quadratic program, a feasible solution  $x^* \in R^n$  will be a strict optimal solution to the problem if and only if  $d^T \nabla^2 f(x^*) d > 0$  for all  $d \in T(x^*)$ ,  $d \neq 0$  (86).

Let  $T_0(x^*) = \{d \in R^n : C_0(x^*)d = 0\}$  and  $T^0(x^*) = \{d \in R^n : C^0(x^*)d = 0\}$ , where  $C_0$  and  $C^0$  are as defined in Theorem 6.3.1 and Theorem 6.3.2. Then, clearly,  $T_0(x^*) \subseteq T(x^*) \subseteq T^0(x^*)$ .

In particular, when all the active inequality constraints are strongly active at  $x^*$ ,  $C_0(x^*) = C^0(x^*)$  and  $T_0(x^*) = T^0(x^*)$ . It follows that if the general constrained optimization problem in (6.5) is a quadratic program, then  $x^*$  will be a strict optimal solution to the problem if and only if  $d^T \nabla^2 f(x^*)d > 0$  for all  $d \in T(x^*) = T_0(x^*) = T^0(x^*)$ ,  $d \neq 0$ .

We now consider the GKproblem in (6.2), which is a typical quadratic program and  $\nabla^2 f(x^*) = -A$ . Let  $Z_0$  and  $Z^0$  be the null space matrices of  $C_0(x^*)$  and  $C^0(x^*)$ , respectively. If all the active inequality constraints are strongly active at  $x^*$ ,  $C_0(x^*) = C^0(x^*)$ ,  $T_0(x^*) = T^0(x^*)$ , and  $Z_0 = Z^0$ . Let  $Z = Z_0 = Z^0$ . Then,  $Z \in R^{n \times (n-m)}$ , and  $T(x^*) = T_0(x^*) = T^0(x^*) = \{d \in R^n : d = Zy : \forall y \in R^{n-m}\}$ , where  $m$  is the row rank of  $C_0(x^*)$  and  $C^0(x^*)$ . It follows that  $x^* \in S$  is a strict optimal solution to the problem if and only if  $y^T Z^T A Z y < 0$  for all  $y \in R^{n-m}$ ,  $y \neq 0$ . More accurately, we have

**Theorem 6.3.5.** *Let  $x^* \in S$  be a KKT point for the GKproblem in (6.2). Assume that the active inequalities in  $S$  are all strongly active at  $x^*$ . Then,  $x^* \in S$  is a strict optimal solution to the GKproblem in (6.2) if and only if*

$$y^T Z^T A Z y < 0 \text{ for all } y \in R^{n-m}, y \neq 0, \quad (6.16)$$

i.e., the reduced Hessian of the objective function of the GKproblem in (6.2) at  $x^*$ ,  $Z^T A Z$ , is negative definite.

The second-order optimality conditions presented in this section can be useful for checking the optimality of the solutions for the GKproblems and hence the strategies for the SEgames beyond the conditions given in Theorem 6.2.1 and 6.2.3. In order to apply these conditions, all we need to do is to find the null space matrices  $Z_0$  or  $Z^0$  and the eigenvalues of the reduced Hessians  $Z_0^T A Z_0$  or  $Z^0 T A Z^0$  to see if they are negative semi-definite or negative definite. For example, suppose that we have a KKT point  $x^* \in S$  for the GKproblem in (6.2) at which the only active constraint is the equality constraint  $1 - \sum_i x_i = 0$ . Then,  $C_0(x^*) = C^0(x^*) = \{-e^T\}$ , for which we can construct a null space matrix  $Z = Z_0 = Z^0 \in R^{n \times (n-1)}$  such that  $Z_{i,j} = 0$  for all  $i$  and  $j$ , except for  $Z_{i,i} = 1$  and  $Z_{i+1,i} = -1$ . Then the optimality of  $x^*$  can be tested by checking the eigenvalues of the reduced Hessian  $Z^T A Z$ . If any of the eigenvalues is positive,  $x^*$  is not optimal, and if all the eigenvalues are negative,  $x^*$  must be optimal and even strictly optimal. Here, in both cases,  $x^*$  remains to be an optimal strategy for

the corresponding SEgame in (6.1). However, the stability of the solution may be different, as we will discuss in greater detail in next section.

Note that the second order necessary and sufficient conditions for the optimal solutions of the GKproblem in (6.2) have been discussed in great detail in (85), where, related to our discussion, there are two necessary and sufficient conditions: (1) A feasible solution  $x^* \in S$  for the GKproblem in (6.2) is a strict optimal solution if and only if  $d^T Ad < 0$  for all  $d \in T(x^*)$ ,  $d \neq 0$ , where  $T(x^*)$  is the reduced tangent cone of the problem at  $x^*$ . (2) If all active inequalities for the GKproblem in (6.2) are strongly active at  $x^*$ , then  $x^*$  is a strict optimal solution if and only if  $Z^T AZ$  is negative definite, when  $T(x^*)$  becomes a linear space defined by matrix  $Z$ . In our analysis, corresponding to (1), we have a necessary condition in Theorem 6.3.3 and sufficient condition in Theorem 6.3.4 separately. They are not equivalent to, but are in fact weaker than the condition in (1). The reason for doing so is that the condition in (1) is hard to test. It is equivalent to solving a matrix co-positivity problem, which is NP-hard in general (95). On the other hand, the condition in Theorem 6.3.3 is equivalent to  $d^T Ad < 0$  for all  $d \in T_0(x^*)$ , which is a smaller cone than  $T(x^*)$ , and is actually a linear space defined by  $Z_0$ . Therefore, the condition is equivalent to  $Z_0^T AZ_0$  negative definite, which can be verified in polynomial time (101). Likewise, the condition in Theorem 6.3.4 is equivalent to  $d^T Ad < 0$  for all  $d \in T^0(x^*)$ , which is a larger cone than  $T(x^*)$ , and is actually a linear space defined by  $Z^0$ . Therefore, the condition is equivalent to  $Z^{0T} AZ^0$  negative definite, which can again be verified in polynomial time. In our analysis, corresponding to (2), we have an equivalent necessary and sufficient condition in Theorem 6.3.5. They are equivalent because if all active constraints for the GKproblem in (6.2) are strongly active at  $x^*$ ,  $T(x^*) = T_0(x^*) = T^0(x^*)$  and  $Z = Z_0 = Z^0$ . It follows that  $d^T Ad < 0$ , for all  $d \in T(x^*)$ ,  $d \neq 0$  is equivalent to  $Z^T AZ$  negative definite. This condition is polynomial time verifiable. We do not need to modify it. The second order optimality conditions in Theorem 6.3.3, 6.3.4, and 6.3.5 are the basis for the later development of our second order stability conditions in Section 6.5 and computational methods in Section 6.6.

## 6.4 Evolutionarily Stable States

An important concept in evolutionary game theory is the evolutionary stability of an optimal strategy. It characterizes the ability of a population to resist small changes or invasions when at equilibrium. Let  $x^* \in S$  be an optimal strategy. Then, the population is at equilibrium state  $x^*$ . Let  $x \in S$  be another arbitrary strategy. Mix  $x^*$  and  $x' \neq x^*$  so that the population changes to a new state,  $\epsilon x + (1 - \epsilon)x^*$ , for some small fraction  $\epsilon > 0$ . Then,  $x^*$  is said to be evolutionarily stable if it remains as a better response to the new “invaded” population state. More accurately, we have the following definition.

**Definition 6.4.1** ((99; 102)). An optimal strategy  $x^* \in S$  for an evolutionary game defined by a fitness matrix  $A$  is evolutionarily stable if there is a small number  $\epsilon' \in (0, 1)$  such that for any  $x \in S$ ,  $x \neq x^*$ ,

$$x^{*T} A(\epsilon x + (1 - \epsilon)x^*) > x^T A(\epsilon x + (1 - \epsilon)x^*), \quad 0 < \epsilon \leq \epsilon'. \quad (6.17)$$

Usually, it is not easy to prove the evolutionary stability of the optimal strategies for an evolutionary game based on its definition. A more straightforward condition is to consider the strategies  $y$  in a small neighborhood  $U$  of the optimal strategy  $x^*$  and check if no  $y \neq x^*$  prevails  $x^*$  such that  $y^T A y \geq x^{*T} A y$ . It turns out that this condition is necessary and also sufficient:

**Theorem 6.4.2** ((99; 102)). *An optimal strategy  $x^* \in S$  for an evolutionary game is evolutionarily stable if and only if there is a small neighborhood  $U$  of  $x^*$  such that*

$$y^T A y < x^{*T} A y \text{ for all } y \in U \cap S, y \neq x^*. \quad (6.18)$$

Note that a SEgame is an evolutionary game. Therefore, the condition in (6.18) also applies to a SEgame. For a SEgame,  $x^{*T} A y = y^T A x^*$  since  $A$  is symmetric. Then,  $y^T A y < x^{*T} A x^*$  for all  $y \in U \cap S$ ,  $y \neq x^*$  since  $y^T A x^* \leq x^{*T} A x^*$  for all  $y \in S$ . This implies that if  $x^*$  is an evolutionary stable strategy for a SEgame, it must be a strict local maximizer of the corresponding GKproblem. It turns out that the converse is also true. We state this property in the following theorem, and also provide a slightly different proof from those given in (99; 102).

**Theorem 6.4.3** ((99; 102)). *An optimal strategy  $x^* \in S$  for a SEgame in (6.1) is evolutionarily stable if and only if it is a strict local maximizer of the corresponding GKproblem in (6.2).*

*Proof.* Let  $x^* \in S$  be an evolutionarily stable strategy for the SEgame in (6.1). Then, the necessary condition follows directly from Theorem 6.4.2, as we have discussed above.

To prove the sufficiency, we assume that  $x^*$  is a strict local maximizer of the GKproblem in (6.2). Then, there must be a neighborhood  $U = \{y \in R^n : \|y - x^*\| < \epsilon' < 2\}$  of  $x^*$  such that for any  $y \in U \cap S$ ,  $y \neq x^*$ ,  $y^T A y < x^{*T} A x^*$ . Let  $x \in S$  be any mixed strategy. Let  $y = \epsilon x + (1 - \epsilon)x^*$ ,  $0 < \epsilon < 1$ . Note that  $\|x - x^*\| \leq \|x\| + \|x^*\| < 2$ , and  $\|y - x^*\| = \epsilon\|x - x^*\| < 2\epsilon$ . Then, for all  $\epsilon < \epsilon'/2 < 1$ ,  $y \in U$  and  $y^T A y < x^{*T} A x^*$ . Note also that

$$\begin{aligned} y^T A y &= (\epsilon x + (1 - \epsilon)x^*)^T A (\epsilon x + (1 - \epsilon)x^*) \\ &= (x^* + \epsilon(x - x^*))^T A (x^* + \epsilon(x - x^*)) \\ &= x^{*T} A x^* + 2\epsilon(x - x^*)^T A (x^* + \epsilon(x - x^*)/2). \end{aligned}$$

It follows that

$$(x - x^*)^T A (x^* + \epsilon(x - x^*)/2) < 0 \text{ for all } \epsilon < \epsilon'/2 < 1.$$

Replace  $\epsilon/2$  by  $\epsilon$  and  $\epsilon'/4$  by  $\epsilon'$ . Then,

$$\begin{aligned} (x - x^*)^T A (x^* + \epsilon(x - x^*)) &< 0 \text{ for all } \epsilon < \epsilon' < 1, \text{ i.e.,} \\ x^T A (x^* + \epsilon(x - x^*)) &< x^{*T} A (x^* + \epsilon(x - x^*)) \text{ for all } \epsilon < \epsilon' < 1. \end{aligned}$$

Since the above inequality holds for all  $x \in S$ , by Definition 6.4.1,  $x^*$  must be an evolutionarily stable strategy for the SEgame in (6.1).  $\square$

## 6.5 Second-order Stability Conditions

By combining Theorem 6.4.3 with the second-order optimality conditions for the optimal solutions to the GKproblem in (6.2) derived in Section 6.3, we can easily obtain a set of second-order stability conditions for the optimal strategies for the SEgame in (6.1): Let  $x^* \in S$  be an optimal strategy for the SEgame in (6.1). Let  $C_0(x^*)$  be a matrix with  $\{e_i^T : x_i^* = 0\}$  and  $\{-e^T\}$  being the rows, where  $e_i$  is the  $i$ th unit vector and  $e = \sum_i e_i$ .

**Theorem 6.5.1.** *Let  $x^* \in S$  be an evolutionarily stable strategy for the SEgame in (6.1). Let the row rank of  $C_0(x^*)$  be equal to  $m$ . Let  $Z_0 \in R^{n \times (n-m)}$  be the null space matrix of  $C_0(x^*)$ . Then,  $Z_0^T A Z_0$  must be negative semi-definite.*

*Proof.* If  $x^* \in S$  is an evolutionarily stable strategy for the SEgame in (6.1), then by Theorem 6.4.3, it must be a strict local maximizer of the GKproblem in (6.2). It follows from Theorem 6.3.3 that  $Z_0^T A Z_0$  must be negative semi-definite.  $\square$

Now, let  $x^* \in S$  be an optimal strategy for the SEgame in (6.1). Let  $C^0(x^*)$  be a matrix with  $\{e_i^T : x_i^* = 0 \text{ and } \mu_i^* > 0\}$  and  $\{-e^T\}$  being the rows, where  $e_i$  is the  $i$ th unit vector,  $e = \sum_i e_i$ , and  $\mu_i^* = x^{*T} A x^* - (A x^*)_i$ .

**Theorem 6.5.2.** *Let  $x^* \in S$  be an optimal strategy for the SEgame in (6.1). Let the row rank of  $C^0(x^*)$  be equal to  $m$ . Let  $Z^0 \in R^{n \times (n-m)}$  be the null space matrix of  $C^0(x^*)$ . If  $Z^{0T} A Z^0$  is negative definite, then  $x^*$  must be an evolutionarily stable strategy.*

*Proof.* If  $x^* \in S$  is an optimal strategy for the SEgame in (6.1), then by Corollary 6.2.4, it must be a KKT point for the GKproblem in (6.2). Therefore, if  $Z^{0T} A Z^0$  is negative definite,  $x^*$  must be a strict local maximizer of the GKproblem in (6.2) by Theorem 6.3.4 and an evolutionarily stable strategy for the SEgame in (6.1) by Theorem 6.4.3.  $\square$

Finally, let  $x^* \in S$  be an optimal strategy for the SEgame in (6.1). If  $\mu_i^* > 0$  for all  $i$  such that  $x_i^* = 0$ , i.e., all the active inequalities in  $S$  are strongly active at  $x^*$ , then  $C_0(x^*) = C^0(x^*)$ , and  $Z_0 = Z^0$ . Let  $Z = Z_0 = Z^0$ .

**Theorem 6.5.3.** *Let  $x^* \in S$  be an optimal strategy for the SEgame in (6.1). Assume that the active inequalities in  $S$  are all strongly active at  $x^*$ . Then,  $x^* \in S$  is an evolutionarily stable strategy for the SEgame in (6.1) if and only if  $Z^T A Z$  is negative definite.*

*Proof.* If  $x^* \in S$  is an optimal strategy for the SEgame in (6.1), then by Corollary 6.2.4, it must be a KKT point for the GKproblem in (6.2). Therefore,  $x^*$  is a strict local maximizer of the GKproblem in (6.2) if and only if  $Z^T A Z$  is negative definite by Theorem 6.3.5 and an evolutionarily stable strategy for the SEgame in (6.1) by Theorem 6.4.3.  $\square$

Although Theorem 6.5.1, 6.5.2, and 6.5.3 are simple extensions from Theorem 6.3.3, 6.3.4, and 6.3.5, they have great implications in practice, for they can be used to check the evolutionary stability of the optimal strategies for the SEgame in (6.1) directly. For example, if the fitness matrix  $A$  is positive

definite, the reduced Hessian  $Z_0^T AZ_0$  will never be negative semi-definite unless the dimension of the null space of  $C_0(x^*)$  is zero or in other words, unless the row rank of  $C_0(x^*)$  is  $n$ . Then,  $x_i^* = 0$  for all but one  $i$ , and the optimal and stable strategies of the SEgame in (6.1) can only be pure strategies. On the other hand, if the fitness matrix  $A$  is negative definite, the reduced Hessian  $Z^{0T} AZ^0$  will always be negative definite unless the dimension of the null space of  $C^0(x^*)$  is zero, and then, all optimal and non-pure strategies for the SEgame in (6.1) will be evolutionarily stable. Even when  $C^0(x^*)$  is only of rank one, i.e.,  $\sum_i x_i^* = 1$  but  $x_i^* > 0$  for all  $i$ ,  $x^*$  is still evolutionarily stable.

Note that an optimal strategy for the SEgame in (6.1) must be a KKT point of the GKproblem in (6.2), but it may not be a local maximizer of the GKproblem in (6.2). It could be a local minimizer or saddle point for the GKproblem in (6.2). Even if it is a local maximizer of the GKproblem in (6.2), it may not be evolutionary stable unless it is a strict local maximizer of the GKproblem in (6.2). In other words, as a KKT point for the GKproblem in (6.2), an optimal strategy for the SEgame in (6.1) could be a local maximizer, local minimizer, or saddle point of the GKproblem in (6.2) while evolutionarily unstable.

Since the second-order stability conditions in Theorem 6.5.1 and 6.5.2 are derived from Theorem 6.3.3 and 6.3.4, they are in different but weaker forms from those given in (85) as well. As we have mentioned in the end of Section 6.3, the advantage of introducing these forms is that they can be checked more efficiently in polynomial time than that given in (85). The latter is equivalent to a matrix co-positivity problem and can be NP-hard to compute. The condition in Theorem 6.5.3 is equivalent to the one given in (85) since it can be verified in polynomial time as those in Theorem 6.5.1 and 6.5.2.

## 6.6 Computational Methods

As we have discussed in previous sections, in order to test the second-order optimality or stability conditions, all we need to do is to form a reduced Hessian for the objective function of the GKproblem in (6.2) and see if it is negative semidefinite or negative definite. The Hessian of the objective function of the GKproblem in (6.2) is basically the fitness matrix  $A$ , while the reduced Hessian is  $Z_0^T AZ_0$  or  $Z^{0T} AZ^0$ , where  $Z_0$  and  $Z^0$  are the null space matrices of  $C_0(x^*)$  and  $C^0(x^*)$ , respectively, for  $x^* \in S$  to be tested,  $C_0(x^*) = \{e_i^T : x_i^* = 0\} \cup \{-e^T\}$  and  $C^0(x^*) = \{e_i^T : x_i^* = 0 \text{ and } \mu_i^* > 0\} \cup \{-e^T\}$ .

There are three major steps to complete a second-order optimality or stability test: (1) Compute the null space matrices  $Z_0$  or  $Z^0$ . (2) Form the reduced Hessians  $Z_0^T A Z_0$  or  $Z^{0T} A Z^0$ . (3) Compute the eigenvalues of the reduced Hessians. In step (1), it can be computationally expensive to find the null space matrix for a given matrix using a general approach, say the QR factorization, which typically requires  $O((n - m)n^2)$  floating-point calculations (101) if  $Z_0$  or  $Z^0$  is a  $n \times (n - m)$  matrix. In step (2), each of the reduced Hessians involves two matrix-matrix multiplications, which also requires  $O(2(n - m)n^2)$  floating-point calculations. However, because of the special structures of  $C_0(x^*)$  and  $C^0(x^*)$ , the calculations in step (1) and step (2) can actually be carried out in a very simple way, without much computational cost:

First of all, the matrices  $C_0(x^*)$  and  $C^0(x^*)$  do not need any computation. They can be constructed straightforwardly as follows: First, form an  $(n + 1) \times n$  matrix with the  $i$ th row equal to  $e_i^T$  and the last row equal to  $-e^T$ , where  $e_i$  is the  $i$ th unit vector and  $e = \sum_i e_i$ . Then, for  $C_0(x^*)$ , remove row  $i$  such that  $x_i^* > 0$ ; for  $C^0(x^*)$ , in addition to row  $i$  such that  $x_i^* > 0$ , remove row  $i$  such that  $x_i^* = 0$  and  $\mu_i^* = 0$ . We demonstrate the structure of  $C_0(x^*)$  and  $C^0(x^*)$  in the following matrix form:

$$C_0(x^*) = \begin{pmatrix} & & & & \cdots \\ & & & & \cdots \\ 0 & \cdots & 1 & \cdots & 0 \\ & & \cdots & & \\ & & \cdots & & \\ -1 & \cdots & -1 & \cdots & -1 \end{pmatrix} \Leftrightarrow e_i^T \text{ such that } x_i^* = 0$$

$$C^0(x^*) = \begin{pmatrix} & & & & \cdots \\ & & & & \cdots \\ 0 & \cdots & 1 & \cdots & 0 \\ & & \cdots & & \\ & & \cdots & & \\ -1 & \cdots & -1 & \cdots & -1 \end{pmatrix} \Leftrightarrow e_i^T \text{ such that } x_i^* = 0 \text{ and } \mu_i^* > 0$$

Next, given the simple structure of  $C_0(x^*)$  and  $C^0(x^*)$ , we in fact do not have to compute the null space matrices  $Z_0$  and  $Z^0$ , either. They can also be constructed easily: First, form an  $n \times n$  identity

matrix with row  $k$  replaced by  $-e^T$  for some  $k$  such that  $x_k^* > 0$ . Then, remove the  $k$ th column; in addition, for  $Z_0$ , also remove column  $j$  such that  $x_j^* = 0$ ; for  $Z^0$ , only remove column  $j$  such that  $x_j^* = 0$  and  $\mu_j^* > 0$ . The following are the matrix forms of  $Z_0$  and  $Z^0$ :

$$Z_0 = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ & \cdots & & & \\ -1 & \cdots & -1 & \cdots & -1 \\ & \cdots & & & \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \Leftarrow \text{row } k \text{ such that } x_k^* > 0 \text{ for some } k$$

(Remove column  $k$ . In addition, also remove column  $j$  such that  $x_j^* = 0$ .)

$$Z^0 = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ & \cdots & & & \\ -1 & \cdots & -1 & \cdots & -1 \\ & \cdots & & & \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \Leftarrow \text{row } k \text{ such that } x_k^* > 0 \text{ for some } k$$

(Remove column  $k$ . In addition, remove only column  $j$  such that  $x_j^* = 0$  and  $\mu_j^* > 0$ .)

It is easy to see that  $Z_0$  or  $Z^0$  are of full column rank  $n - m$ , where  $m$  is the row rank of  $C_0(x^*)$  or  $C^0(x^*)$ . It is also easy to verify that  $C_0(x^*)Z_0 = 0$  and  $C^0(x^*)Z^0 = 0$ , and therefore,  $Z_0$  and  $Z^0$  can indeed be used as null space matrices of  $C_0(x^*)$  and  $C^0(x^*)$ , respectively. Yet, the construction of  $Z_0$  and  $Z^0$  does not have computational cost at all.

Finally, with  $Z_0$  and  $Z^0$  as given above, the computation of the reduced Hessians  $Z_0^T AZ_0$  or  $Z^{0T} AZ^0$  does not require full matrix-matrix multiplications. Let  $H = Z^T AZ$  with  $Z = Z_0$  or  $Z^0$ . We show how  $H$  can be calculated with less computational cost: Let  $B = AZ$ . Then,  $H = Z^T AZ = Z^T B$ . Let  $B^j$  and  $Z^j$  be column  $j$  of  $B$  and  $Z$ , respectively. Assume that  $Z^j = e_i - e_k$  for some  $i$ . Then,  $B^j = AZ^j$  can be obtained by subtracting column  $k$  from column  $i$  of  $A$  with  $n$  floating-point calculations. Since  $B$  has only  $n - m$  columns, the computation of  $B$  requires  $n(n - m)$  floating-point calculations. Let  $H_i$  and  $Z^{iT}$  be row  $i$  of  $H$  and  $Z^T$ . Also assume that  $Z^{iT} = e_j^T - e_k^T$  for some  $j$ . Then,  $H_i = Z^{iT} B$  can be obtained by subtracting row  $k$  from row  $j$  of  $B$  with  $n - m$  floating-point calculations. Since  $H$  has only  $n - m$  rows, the computation of  $H$  requires  $(n - m)^2$  floating-point

calculations. By putting the calculations for  $B$  and  $H$  together, we then obtain the computation for the whole reduced Hessian  $Z^T AZ$  to be  $(n - m)(2n - m)$  floating-point calculations, which is much less costly than full matrix-matrix multiplications.

## 6.7 Games for Genetic Selection

A genetic selection problem and in particular, the problem for allele selection at single or multiple genetic loci can be formulated as a symmetric evolutionary game. Recall that the fitness of different allele pairs or in other words, different genotypes at a given genetic locus can be given in a matrix with the rows corresponding to the choices for the first allele and the columns to the choices for the second allele in the allele pairs. If there are  $n$  different alleles, there will be  $n$  different choices for both the first and second alleles, and the fitness matrix will be an  $n \times n$  matrix. With such a fitness matrix, the allele selection game can be defined with the choices of the first and second alleles as the strategies for player I and player II of the game, where player I can be considered as a specific individual and player II as a typical individual in the given population. If there are  $n$  different alleles, the strategy for player I can be represented by a vector  $x \in R^n$ ,  $x \geq 0$ ,  $\sum_i x_i = 1$ , and the strategy for player II by a vector  $y \in R^n$ ,  $y \geq 0$ ,  $\sum_i y_i = 1$ . Let the fitness matrix be given by  $A \in R^{n \times n}$ . Let  $S = \{x \in R^n : x \geq 0, \sum_i x_i = 1\}$ . The average fitness of an allele choice  $x \in S$  in an allele population  $y \in S$  will be  $x^T A y$ . We then want to find an optimal choice of  $x^* \in S$  such that

$$x^{*T} A x^* \geq x^T A x^* \text{ for all } x \in S, \quad (6.19)$$

i.e., in allele population  $x^*$ , any individual with allele choice  $x$  other than  $x^*$  will not have a better average fitness than allele choice  $x^*$ . Note that the fitness for allele pair  $(i, j)$  usually is the same as that for  $(j, i)$ . Therefore, the fitness matrix for allele selection is typically symmetric, and the game in (6.19) is then a SEgame.

As we have discussed in previous sections, the selection game in (6.19) can be studied with a generalized knapsack problem:

$$\max_{x \in R^n} x^T A x / 2 \quad (6.20)$$

$$\text{subject to } \sum_i x_i = 1, x \geq 0.$$

By Corollary 6.2.4, an optimal strategy of the selection game in (6.19) is equivalent to a KKT point of the GKproblem in (6.20), and by Theorem 6.4.3, if it is evolutionarily stable, it must correspond to a strict local maximizer of the GKproblem in (6.20), and vice versa. In addition, the optimality and stability conditions derived in previous sections all apply to the selection game in (6.19). We demonstrate the applications of these conditions with several example selection games including some from the study of genetic disorders.

We first consider a genetic locus with two alleles, one dominant and another recessive. Many genetic traits are due to the genotypic differences in a specific locus of two alleles. For example, in the well-known Mendel's experiment, the color of the flowers depends on the pair of alleles at certain genetic locus, one for pink color and dominant, and another for white and recessive. Let the dominant allele be denoted by  $A$  and recessive one by  $a$ . There can be four possible allele pairs,  $AA$ ,  $Aa$ ,  $aA$ , and  $aa$ . Since  $A$  is dominant,  $AA$ ,  $Aa$ , and  $aA$  will produce pink flowers, while  $aa$  will produce white ones (see Fig. 6.2). According to the Hardy-Weinberg Law, if pink flowers and white flowers have the same selection chance, the distributions of the genotypes  $AA$ ,  $Aa$ ,  $aA$ , and  $aa$  and the alleles  $A$  and  $a$  in the population will not change over generations. Otherwise, different genotypes may have different fitness, and some may be selected while others eliminated (88).

Indeed, some alleles, either dominant or recessive, may cause genetic disorders. When they are dominant, both homozygote and heterozygote pairs containing the dominant allele will cause the disorders. When they are recessive, only the homozygote pairs of two recessive alleles will cause the problem. In either case, the genotypes that cause the genetic disorders will have lower fitness than those that do not. For example, cystic fibrosis is a disease caused by a recessive allele. The normal allele or the dominant one codes for a membrane protein that supports the transportation of ions for cells. It functions normally even when in the heterozygote form with one abnormal allele. However, if both alleles are the recessive ones, there will not be normal membrane protein expressions, giving rise to the cystic fibrosis disease. A further example is the Huntington's disease, a degenerative disease of the nerve system, caused by a lethal dominant allele. Both homozygote and heterozygote pairs of alleles containing the dominant allele will be harmful. Only the homozygote pairs of the recessive alleles will be normal. Fortunately, this harmful allele is rare in human population (88).

We now use a  $2 \times 2$  matrix  $A$  to represent the fitness of the four possible pairs of alleles, with  $A_{1,1}$  for  $AA$ ,  $A_{1,2}$  for  $Aa$ ,  $A_{2,1}$  for  $aA$ , and  $A_{2,2}$  for  $aa$ . We assume that the fitness values for  $Aa$  and  $aA$  are the same. Therefore,  $A_{1,2} = A_{2,1}$ , and  $A$  is a symmetric matrix. Let  $x = (x_1, x_2)^T$  be the strategy for choosing between allele  $A$  and  $a$ . We then have a SEgame for this allele selection problem, to find an optimal strategy  $x^* \in S$  such that

$$x^{*T} Ax^* \geq x^T Ax^* \text{ for all } x \in S, \quad (6.21)$$

where  $S = \{x \in R^2 : x \geq 0, \sum_i x_i = 1\}$ . We have also a corresponding GKproblem:

$$\max_{x \in R^2} x^T Ax / 2 \quad (6.22)$$

$$\text{subject to } \sum_i x_i = 1, x \geq 0.$$

We analyze the solutions and their stabilities of this game for the following three cases:

**Case 1.**  $A_{1,1} = A_{1,2} = A_{2,1} > A_{2,2}$ : Here, the fitness of the homozygote pair  $AA$  is the same as that of the heterozygote pairs  $Aa$  and  $aA$ , but the fitness of the homozygote pair  $aa$  is lower than all other cases. This is a case such as that for the cystic fibrosis disease, when the recessive gene causes illness. In this case, the SEgame in (6.21) has a single solution  $x^* = (1, 0)^T$ , and  $x^{*T} Ax^* = A_{1,1}$ . This implies that the population will eventually reach an equilibrium state with only the dominant gene  $A$  left. Let  $x \in S$  be any other possible strategy. It is easy to verify that

$$x^T Ax = A_{1,1}x_1^2 + A_{1,2}x_1x_2 + A_{2,1}x_1x_2 + A_{2,2}x_2^2 < A_{1,1} = x^{*T} Ax^*,$$

for all  $x \neq x^*$ , i.e.,  $x_2 > 0$ . Therefore,  $x^*$  is a strict local maximizer of the GKproblem in (6.22), and by Theorem 6.4.3, it is an evolutionarily stable state.

**Case 2.**  $A_{1,1} = A_{1,2} = A_{2,1} < A_{2,2}$ : Here, the fitness of the homozygote pair  $AA$  is equal to that of the heterozygote pairs  $Aa$  and  $aA$ , but lower than that of the homozygote pair  $aa$ . This is a case such as that for the Huntington's disease, when the dominant gene causes illness. In this case, the SEgame in (6.21) has two solutions:  $x^* = (1, 0)^T$  with  $x^{*T} Ax^* = A_{1,1}$  and  $x^* = (0, 1)^T$  with  $x^{*T} Ax^* = A_{2,2}$ . The first solution implies that the population may eventually reach an equilibrium state with only the dominant gene  $A$  left. The second solution implies that the population may eventually reach an equilibrium state with only the recessive gene  $a$  left. Let  $x \in S$  be any other possible strategy. For the first solution

$x^* = (1, 0)^T$ , It is easy to verify that

$$x^T Ax = A_{1,1}x_1^2 + A_{1,2}x_1x_2 + A_{2,1}x_1x_2 + A_{2,2}x_2^2 > A_{1,1} = x^{*T} Ax^*,$$

for all  $x \neq x^*$ , i.e,  $x_2 > 0$ . Therefore,  $x^*$  is not even a local maximizer of the GKproblem in (6.22), and by Theorem 6.4.3, it is not evolutionarily stable. For the second solution  $x^* = (0, 1)^T$ , It is easy to verify that

$$x^T Ax = A_{1,1}x_1^2 + A_{1,2}x_1x_2 + A_{2,1}x_1x_2 + A_{2,2}x_2^2 < A_{2,2} = x^{*T} Ax^*,$$

for all  $x \neq x^*$ , i.e,  $x_1 > 0$ . Therefore,  $x^*$  is a strict local maximizer of the GKproblem in (6.22), and by Theorem 6.4.3, it is an evolutionarily stable state.

**Case 3.**  $A_{1,2} = A_{2,1} > A_{1,1} > A_{2,2}$ : Here, the fitness of the homozygote pairs  $AA$  and  $aa$  are lower than that of the heterozygote pairs  $Aa$  and  $aA$ . Such a case could happen for example in the study of malaria infection, where  $A$  represents the wild-type gene, while  $a$  represents the mutated gene. Individuals with  $AA$  types are susceptible to malaria infection, while those with  $Aa$  and  $aA$  types appear to be able to resist the infection. However, when  $aa$  types are formed, the individuals will develop a serious disease called the sickle cell disease. In any case, the SEgame in (6.21) has a single solution  $x^* = (x_1^*, x_2^*)$ ,  $x_1^* > 0$ ,  $x_2^* > 0$ , and

$$x_1^* = (A_{1,2} - A_{2,2})/(A_{1,2} + A_{2,1} - A_{1,1} - A_{2,2})$$

$$x_2^* = (A_{1,1} - A_{1,2})/(A_{1,2} + A_{2,1} - A_{1,1} - A_{2,2})$$

Since both  $x_1^* > 0$  and  $x_2^* > 0$ , it is easy to construct a null space matrix  $Z = (1, -1)^T$ , and see that  $Z^T AZ = A_{1,1} + A_{2,2} - A_{1,2} - A_{2,1} < 0$ . Therefore, by Theorem 6.3.5,  $x^*$  must be a strict local maximizer of the GKproblem in (6.22), and by Theorem 6.4.3 or 6.5.3, it is an evolutionarily stable state.

Next, we consider a more complicated case related to genetic mutations for malaria infections. In Africa and Southeast Asia, where human population has been exposed to serious malaria infection, certain genetic mutations have survived for a gene that codes the hemoglobin proteins of blood cells. These mutations resist malaria infection, but may cause other serious illness as well when in homozygote forms such as the sickle cell disease. Here we consider three well-studied allele forms of this gene,

the wild type,  $S$ -mutation, and  $C$ -mutation, denoted by  $W$ ,  $S$ , and  $C$  alleles. The normal genotype would be  $WW$ , but subnormal ones include  $WS$ ,  $WC$ , and  $SC$ , which may have malaria resistance functions. Other forms,  $SS$  and  $CC$ , may cause other illness. These functions can be described with a  $3 \times 3$  fitness matrix  $A$ , with rows corresponding to the choices of  $W$ ,  $S$ , and  $C$  for the first allele, and the columns to the choices of  $W$ ,  $S$ , and  $C$  for the second allele, when forming the allele pairs or in other words, the genotypes. Based on an estimate given in (100), this fitness matrix can be defined as follows:

|     | $W$  | $S$  | $C$  |
|-----|------|------|------|
| $W$ | 0.89 | 1.00 | 0.89 |
| $S$ | 1.00 | 0.20 | 0.70 |
| $C$ | 0.89 | 0.70 | 1.31 |

From this matrix, we see that the genotype  $WS$  has good fitness, while  $CC$  is the best. The genotype  $WW$  is not very good because it is susceptible to malaria infection, while  $SS$  is the worse because it causes the sickle cell disease. We may wonder how the alleles will eventually distribute in the population under such selection pressures. We have solved a SEgame with this fitness matrix and obtained three solutions:  $x^{(1)} = (0, 0, 1)^T$ ,  $x^{(2)} = (0.879, 0.121, 0)^T$ , and  $x^{(3)} = (0.832, 0.098, 0.070)^T$ . The first solution suggests that the population may end up with all  $C$  alleles since the genotype  $CC$  seems have the best fitness. The second solution suggests a large portion of  $W$  alleles, with a small percentage of  $S$  alleles, which increases the resistance to malaria infection, yet does not have a large chance for  $SS$  combinations. The third solution means that the three alleles may co-exist.

We have also solved a corresponding GKproblem with the above matrix  $A$ , using a Matlab code. It turned out that we have only found two local maximizers for the GKproblem corresponding to  $x^{(1)}$  and  $x^{(2)}$ . At least, computationally, we have not found  $x^{(3)}$  as a local maximizer, which suggests that  $x^{(1)}$  and  $x^{(2)}$  may be evolutionarily stable, while  $x^{(3)}$  may not. Indeed, at solution  $x^{(3)}$ , the only active constraint for the GKproblem is  $\sum_i x_i = 1$ . The null space matrix  $Z$  for the Jacobian of this equation can be constructed as

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix}.$$

We then have the reduced Hessian of the GKproblem to be

$$Z^T AZ = \begin{pmatrix} 0.42 & 0.72 \\ 0.72 & 0.11 \end{pmatrix},$$

and the eigenvalues of this reduced Hessian are  $-0.4715$  and  $1.0015$ . By Theorem 6.3.5,  $x^{(3)}$  is not a local maximizer of the GKproblem and hence by Theorem 6.4.3 or 6.5.3, it is not evolutionarily stable.

Based on the above analysis, we would predict that  $x^{(3)}$  for the co-existing of three alleles in the population will never happen because it is unstable. The solution  $x^{(1)}$  corresponds to a global maximizer of the GKproblem. Based on our simulation (not shown), it also has a large attraction region in the sense that most solutions would converge to  $x^{(1)}$  unless the initial value for  $C$  allele is very small, say less than 5%. In current population,  $C$  allele is indeed rare and therefore, the population does not have much chance to evolve to this state. The population have typically a large percentage of  $W$  alleles, a small percentage of  $S$  alleles, and some rare  $C$  alleles, and therefore, we would predict that  $x^{(2)}$  will be the most likely and stable state of the population in the end.

## 6.8 Concluding Remarks

In this paper, we have reviewed the theory for obtaining optimal and stable strategies for SEgames, and provided some new proofs and computational methods. In particular, we have reviewed the relationship between the SEgame and the GKproblem, and discussed the first and second order necessary and sufficient conditions that can be derived from this relationship for testing the optimality and stability of the strategies. Some of the conditions are given in different forms from those in previous work and can be verified more efficiently. We have also derived more efficient computational methods for the evaluation of the conditions than conventional approaches. We have demonstrated how these conditions can be applied to justifying the strategies and their stabilities for a special class of genetic selection games including some in the study of genetic disorders. Further studies can be pursued in the following possible directions though:

First, novel methods can be developed for solving special types of SEgames and especially for obtaining the evolutionarily stable strategies for the games by solving some special classes of GKproblems.

For example, if the fitness matrix for a SEgame is negative definite, then the corresponding GKproblem is a strictly convex quadratic program and can be solved efficiently using some special algorithms (87). Further, the solution is guaranteed to be a strict local maximizer for the GKproblem and hence an evolutionarily stable strategy for the SEgame. A more complicated case is when the fitness matrix is positive definite. Then, only pure strategies may be evolutionarily stable. A special algorithm can then be developed to only find the solutions for the GKproblem that correspond to the pure strategies of the SEgame.

Second, in Theorem 6.3.5 and 6.5.3, we have stated two optimality and stability conditions. They are necessary and sufficient, but require all active constraints to be strongly active at  $x^*$ , when  $C_0(x^*) = C^0(x^*)$ ,  $T_0(x^*) = T^0(x^*)$ , and  $Z_0 = Z^0$ . However, in practice, this assumption may not hold. A more general necessary and sufficient condition, without the above assumption, is to require  $d^T A d < 0$  for all  $d \in T(x^*)$ ,  $d \neq 0$ , where  $T(x^*)$  is the reduced tangent cone at  $x^*$ , as given in (85). As we have mentioned in previous sections, this condition is not easy to test. It is equivalent to testing the copositivity of a matrix, which is difficult in general (84; 95). But still, an efficient algorithm may be developed for SEgames and GKproblems for small sizes of problems or problems with special structures.

Third, it is not so hard to verify that the GKproblem is NP-hard in general, because the maximum clique problem can be formulated as a GKproblem (94; 98). However, how to extend this result to the SEgame is not so clear, because the SEgame is not exactly equivalent to the GKproblem. Several related questions are asked: is any maximal clique a local maximizer of the GKproblem for the maximum clique problem? If not, what condition is needed? If yes, is it a strict local maximizer? Is the maximum clique a global maximizer? Is it an evolutionarily stable strategy for the corresponding SEgame? We are interested in all these questions and are trying to find their answers.

Fourth, though not equivalent, the correspondence between the SEgame and GKproblem is interesting. A similar relationship may be found between a class of nonlinear games and nonlinear optimization problems. Indeed, we can define an n-strategy two-player game by a fitness function  $x^T \pi(y)$  with  $\pi(y)$  being a nonlinear function. The game then becomes a nonlinear game. If  $\pi(y)$  is a gradient field, i.e., there is a function  $f(y)$  such that  $\nabla f(y) = \pi(y)$ , then, an optimal strategy  $x^* \in S$  such that  $x^{*T} \pi(x^*) \geq x^T \pi(x^*)$  for all  $x \in S$  corresponds to an optimal solution  $x^* \in S$  such that  $f(x^*) \geq f(x)$

for all  $x$  in a small neighborhood of  $x^*$ ,  $x \in S$ . Then, it would be interesting to see what additional relationships between the SEgame and GKproblem can be extended to their nonlinear cases.

Finally, we have demonstrated the applications of SEgames to allele selection at single genetic loci. They can be extended to alleles at multiple genetic loci, if there is no mutation or recombination. In this case, an individual can be identified by a sequence of alleles at the multiple loci. In other words, a selection strategy will be a choice of a specific sequence of alleles. This would certainly increase the strategy space substantially. For example, if there are two loci  $G_1$  and  $G_2$ , with two possible alleles  $A$  and  $a$  for  $G_1$  and two other possible ones  $B$  and  $b$  for  $G_2$ , then there will be four possible sequences of alleles for the two loci:  $AB$ ,  $Ab$ ,  $aB$ ,  $ab$ , each corresponding to one pure strategy. In general, if there are  $m$  loci  $G_i$ ,  $i = 1, \dots, m$ , with  $m_i$  possible alleles for  $G_i$ , then there will be  $n = \prod_{i=1:m} m_i$  possible sequences of alleles. The number of pure strategies and hence the dimension of the game will be  $n$ , which can be a large number. In any case, in practice, mutation and recombination often are not negligible, and therefore, our model must incorporate such effects. The topics could include other so-called linkage disequilibrium factors, but they are all beyond the scope of this paper (100). We will pursue these issues in our future efforts.

### Acknowledgments

This work is partially supported by the NSF/DMS grant DMS0914354. The authors would like to thank Professor Yaxiang Yuan, Professor Yuhong Dai, and Professor Xin Liu of Chinese Academy of Sciences for many insightful discussions in the early stage of this work. The authors would also like to thank Professor Stephen Wright of University of Wisconsin at Madison for taking time to read the early draft of the paper and providing valuable suggestions.

## CHAPTER 7. OPTIMAL STRATEGIES FOR SOCIAL NETWORK GAMES

A paper submitted to *Journal of Optimization Theory and Applications*

Min Wang, Yuanyuan Huang, Yiping Hao, Wen Zhou, Zhijun Wu

### Abstract

Species make social contacts and form social networks. The latter may have great impacts on the evolution of a population, such as preserving certain genetic features, sharing knowledge and information, preventing invasions, etc. In this paper, we show that the evolution of a population over a social network can be modeled as a symmetric evolutionary game. Its equilibrium states can therefore be obtained and analyzed by solving an optimization problem called the generalized knapsack problem. We show that an equilibrium state often corresponds to a social clique, when the population is distributed evenly on the clique. However, an equilibrium state may or may not be evolutionarily stable, whether it is on a clique or not. Only those stable ones may be observable or sustainable in nature. We analyze several different types of equilibrium states and prove a set of conditions for their stabilities. We show in particular that the equilibrium states on cliques are evolutionarily stable except for special circumstances, while non-clique equilibrium states are unstable in general. Therefore, the optimal clique strategies should have an evolutionary advantage over the non-clique ones.

### 7.1 Introduction

Species make social contacts and form social networks. The evolution of a population over a social network can be of great biological or social interest, and be studied as a so-called social network game with the adjacent matrix of the network being the payoff or fitness matrix. Individuals or species in a social network interact with those they connect. They tend to maximize their social activities, which

then drive the population to migrate to nodes or sites with the most possible "social connections". As such, the social network must play an essential role in evolution, for preserving certain genetic features, sharing knowledge and information, or preventing invasions, etc. (109; 112; 117; 119).

In this paper, we show that a social network game is in fact a symmetric evolutionary game or SEgame for short. Its equilibrium states can therefore be obtained and analyzed by solving an optimization problem called the generalized knapsack problem or GKproblem for short. Based on this connection, we show that an equilibrium state often corresponds to a social clique, when the population is distributed evenly on the clique. However, an equilibrium state may or may not be evolutionarily stable, whether it is on a clique or not. Only those stable ones may be observable or sustainable in nature. We analyze several different types of equilibrium states and prove a set of conditions for their stabilities. We show in particular that the equilibrium states on cliques are evolutionarily stable except for special circumstances, while non-clique equilibrium states are unstable in general. Therefore, the optimal clique strategies should have an evolutionary advantage over the non-clique ones. These properties have not been discussed formally in literature before, but can be interesting for the understanding of the behaviors of social networks of either biological, or cultural, or economic types.

## 7.2 Social Network Games

The social activities of a population can be described by a social network, with the nodes representing the social sites and the links the social channels. The population is distributed over the social sites. The individuals at different social sites may be different species, different cultural groups, different races, or different countries. They make social contacts through the social channels. The types of social contacts may include mating, making friends, communication, sharing knowledge, sharing religious beliefs, or trading. The results from socializing over a social network may include increasing social connections, genetic and cultural diversities, group protection, economic welfare, but also spread of rumors or diseases (109; 112; 117; 119) (see Fig. 7.1).

A network can be represented by a graph  $G = (V, E)$  of  $n$  nodes  $V = \{i = 1, \dots, n\}$  and  $m$  edges  $E = \{(i, j) : \text{There is a link between node } i \text{ and } j\}$ . The connectivity of graph  $G$  can be described by

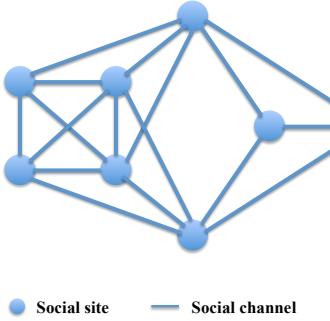


Figure 7.1 Social Networks: A social network consists of social sites and social channels. A population is distributed over the social sites. The individuals at different social sites may be different species, cultural groups, races, or countries. They make social contacts through the social channels. The types of social contacts may be mating, communication, sharing knowledge, or trading. The results from socializing over a social network may include increasing social connections, genetic or cultural diversities, economic welfare, but also spread of rumors or diseases.

a matrix  $A \in R^{n \times n}$ , with  $A_{i,j} = 1$  if  $(i,j) \in E$ , and  $A_{i,j} = 0$  if  $(i,j) \notin E$ ,  $i,j = 1, \dots, n$ . Note that we only consider simple networks, i.e., there is no links connecting a node to itself. Therefore,  $A_{i,i} = 0$  for all  $i = 1, \dots, n$ . The matrix  $A$  is called the adjacency matrix of  $G$  (see Fig. 7.2).

Let  $y \in R^n$  be a population profile, with  $y_j$  being the fraction of species occupying node  $j$ ,  $\sum_j y_j = 1$ . Then,  $y$  shows how the population is distributed over the network. If an individual decides to stay at node  $i$ , the average social contacts he or she can make with the individuals at other nodes would be  $(Ay)_i = \sum_j A_{i,j}y_j$ . Let  $x$  be the strategy of the individual to join the network, with  $x_i$  being the frequency to stay with node  $i$ ,  $\sum_i x_i = 1$ . Then, the average social contacts that this individual can make would be  $x^T Ay = \sum_i x_i (Ay)_i = \sum_{i,j} x_i A_{i,j} y_j$ . We call this value  $x^T Ay$  the social benefit or payoff of individual  $x$  in population  $y$  over network  $G$ . Given population  $y$ , we assume that each individual wants to maximize his or her social payoff by choosing an optimal strategy  $x$ .

Note that the choice of strategy  $x$  for each individual depends on profile  $y$  of the population, but once a decision is made,  $y$  may be changed due to a change in  $x$ . An equilibrium state will be reached in the end, when an optimal strategy  $x^*$  is found such that strategy  $x^*$  prevails in population  $y = x^*$ , i.e.,  $x^{*T} Ax^* \geq x^T Ax^*$  for all strategy  $x$ . Indeed, in such a state, strategy  $x^*$  are optimal and will not change, while the population will remain to be  $x^*$ . As such, the evolution of a population over a social

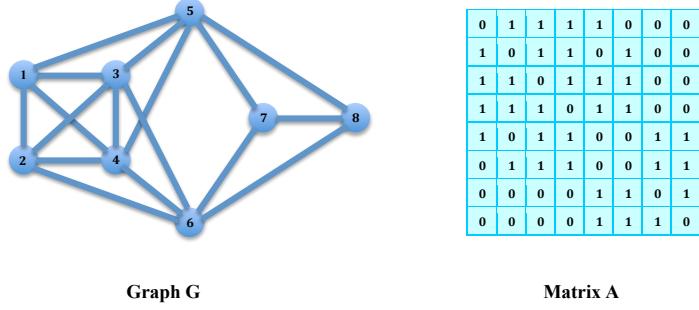


Figure 7.2 Graph Representation: A network can be represented by a graph  $G = (V, E)$  where  $E = \{(i, j) : \text{There is a link between node } i \text{ and } j\}$  and  $V = \{i = 1, \dots, n\}$ . The connectivity of graph  $G$  can be described by a matrix  $A \in R^{n \times n}$ , with  $A_{i,j} = 1$  if  $(i, j) \in E$ , and  $A_{i,j} = 0$  if  $(i, j) \notin E$ ,  $i, j = 1, \dots, n$ . The matrix  $A$  is called the adjacency matrix of  $G$ .

network can be modeled as an evolutionary game (110; 116; 120). We call it a social network game or SNGame for short. We call  $x^*$  an optimal strategy or equilibrium state of the game.

With a goal of maximizing the social activities, a population often migrates to sites with the most possible connections among the sites. Such a group of sites corresponds to a subgraph in the network with a dense set of edges among its nodes. The best one would be a complete subgraph. We call it a social clique. More specifically, we say that a state  $x$  is a social clique if  $x_i > 0$  for all node  $i$  in the clique and  $x_i = 0$  otherwise. In this state, the population is distributed over a clique. Of course, not all the cliques are optimal choices for a game.

### 7.3 Social Network Games as Symmetric Evolutionary Games

Consider a social network game (SNGame) on a network  $G = (V, E)$ ,  $|V| = n$ ,  $|E| = m$ . Let  $S = \{x \in R^n : x \geq 0, \sum_i x_i = 1\}$  be the set of all mixed strategies for the game. Then, the goal of the game is to find an optimal strategy  $x^* \in S$  such that

$$x^{*T} Ax^* \geq x^T Ax^* \text{ for all } x \in S, \quad (7.1)$$

where  $A$  is the adjacency matrix of  $G$  and is used as the payoff matrix of the game. Since  $A$  is always symmetric, the SNGame in (7.1) is a symmetric evolutionary game (SEgame). A general set of opti-

mality and stability conditions for the SNgame in (7.1) can therefore be derived from the theory for the SEgame:

### 7.3.1 Optimality Conditions

For any evolutionary game, it is well known that a mixed strategy  $x^* \in S$  is optimal if and only if it satisfies a set of complementarity conditions:  $x^{*T} Ax^* = (Ax^*)_i$  for all  $i$  such that  $x_i^* > 0$  and  $x^{*T} Ax^* \geq (Ax^*)_i$  for all  $i$  such that  $x_i^* = 0$ . These conditions also apply to any symmetric evolutionary game (SEgame) including the SNgame in (7.1), and can be given more specifically in the following form.

**Theorem 7.3.1.** *A strategy  $x^* \in S$  is an optimal strategy for the SNgame in (7.1) if and only if there is a scalar  $\lambda^*$  such that*

$$x_i^* \geq 0, \lambda^* - (Ax^*)_i \geq 0 \quad (7.2)$$

$$x_i^*(\lambda^* - (Ax^*)_i) = 0, i = 1, \dots, n. \quad (7.3)$$

The proof for the theorem is straightforward and can be found in standard texts (110; 116; 120). Since it is helpful for the understanding of the game, we still provide one here for the self-containedness of the paper:

*Proof.* If  $x^* \in S$  satisfies the conditions in (7.2) and (7.3), by adding all equations in (7.3), we then obtain  $\lambda^* = x^{*T} Ax^*$ . Let  $x \in S$  be an arbitrary strategy. Multiply the second inequality in (7.2) by  $x_i$ . Then, by adding all second inequalities in (7.2), we obtain  $\lambda^* - x^T Ax^* \geq 0$ , i.e.,  $x^{*T} Ax^* \geq x^T Ax^*$ , since  $\lambda^* = x^{*T} Ax^*$ . Therefore,  $x^*$  is an optimal strategy for the SNgame in (7.1).

If  $x^* \in S$  is an optimal strategy for the SNgame in (7.1), then  $x^{*T} Ax^* \geq x^T Ax^*$  for any  $x \in S$  and therefore,  $x^{*T} Ax^* \geq e_i^T Ax^* = (Ax^*)_i$  for all  $i$ . Let  $\lambda^* = x^{*T} Ax^*$ . Then,  $\lambda^* - (Ax^*)_i \geq 0$  for all  $i$ . Assume that  $x_i^*(\lambda^* - (Ax^*)_i) > 0$  for some  $i$ . By adding all the left-hand sides of the equations in (7.3), we then obtain  $\lambda^* > x^{*T} Ax^*$ , which contradicts to the fact that  $\lambda^* = x^{*T} Ax^*$ . Therefore,  $x_i^*(\lambda^* - (Ax^*)_i) = 0$  for all  $i$ .  $\square$

A SEgame is always related to an optimization problem called the generalized knapsack problem (GKproblem) (105; 111; 115). So is a SNgame in (7.1). The GKproblem corresponding to a SNgame

in (7.1) is to find an optimal solution  $x^* \in R^n$  for the following maximization problem:

$$\max_{x \in R^n} x^T Ax/2, \quad (7.4)$$

$$\text{subject to } \sum_i x_i = 1, \quad x \geq 0,$$

where  $A$  is as defined in the SNgame in (7.1). Since the problem is to maximize the social payoff  $x^T Ax$ , we call it the social payoff problem or SPproblem for short.

The SPproblem in (7.4) is a constrained optimization problem. Based on general optimization theory, the optimal solution to a constrained optimization problem must satisfy certain conditions (107; 114). We say that a vector  $x \in R^n$  is a feasible solution for the SPproblem in (7.4) if it satisfies all the constraints of the problem, and that a vector  $x^* \in R^n$  is an optimal solution for the SPproblem in (7.4) if it is feasible for the problem and  $x^{*T} Ax^* \geq x^T Ax$  for all  $x \in R^n$  feasible in a small neighborhood  $U$  of  $x^*$ . Note that  $x$  is feasible for the SPproblem in (7.4) is equivalent to  $x \in S$ , with  $S = \{x \in R^n : x \geq 0, \sum_i x_i = 1\}$  as defined for the SNgame in (7.1). Note also that  $x^*$  is optimal for the SPproblem in (7.4) is equivalent to  $x^*$  is a local maximizer of the problem.

**Theorem 7.3.2.** *If  $x^* \in R^n$  is an optimal solution for the SPproblem in (7.4), then there must be a scalar  $\lambda^*$  such that*

$$x_i^* \geq 0, \quad \lambda^* - (Ax^*)_i \geq 0, \quad (7.5)$$

$$x_i^*(\lambda^* - (Ax^*)_i) = 0, \quad i = 1, \dots, n. \quad (7.6)$$

*Proof.* Note that the Lagrangian function for the SPproblem in (7.4) can be written in the following form:

$$L(x, \lambda, \mu) = -x^T Ax/2 - \lambda(1 - \sum_i x_i) - \mu^T x.$$

where  $x \in R^n$ ,  $\lambda \in R$ ,  $\mu \in R^n$ . Then, by the first-order necessary conditions for constrained optimization, if  $x^* \in R^n$  is an optimal solution for the SPproblem in (7.4), there must be a scalar

$\lambda^* \in R$  and a vector  $\mu^* \in R^n$  such that

$$\begin{aligned}\nabla_x L(x^*, \lambda^*, \mu^*) = -Ax^* + \lambda^*e - \mu^* &= 0, \\ \sum_i x_i^* &= 1, \\ x^* &\geq 0, \\ \mu^* &\geq 0, \\ x^{*T} \mu^* &= 0.\end{aligned}$$

By substituting  $\mu^* = \lambda^*e - Ax^*$  in all the formulas, we then have

$$x^* \geq 0, \lambda^*e - Ax^* \geq 0,$$

$$x^{*T}(\lambda^*e - Ax^*) = 0,$$

which are equivalent to the conditions in (7.5) and (7.6).  $\square$

The first-order necessary conditions for a constrained optimization problem are also called the KKT conditions (107; 114). Vectors that satisfy the KKT conditions are called KKT points. An optimal solution for a constrained optimization problem must be a KKT point, but a KKT point may not necessarily be an optimal solution. Based on the above proof, the conditions in Theorem 7.3.2 are basically the KKT conditions for the SPproblem in (7.4). Therefore, an optimal solution for the SPproblem in (7.4) must be a KKT point for the problem, but a KKT point that satisfy those conditions may not necessarily be an optimal solution.

Notice that the conditions in (7.2) and (7.3) are the same as those in (7.5) and (7.6). However, it does not imply that the SNgame in (7.1) is equivalent to the SPproblem in (7.4), because they are necessary and sufficient for a strategy to be optimal for the SNgame in (7.1) but only necessary for a solution to be optimal for the SPproblem in (7.4). We give a clear distinction between the two problems in the following corollary.

**Corollary 7.3.3.** *An optimal solution  $x^* \in S$  for the SPproblem in (7.4) must be an optimal strategy for the SNgame in (7.1), while an optimal strategy  $x^* \in S$  for the SNgame in (7.1) is only a KKT point for the SPproblem in (7.4), which is necessary but not sufficient to be optimal for the SPproblem in (7.4).*

### 7.3.2 Stability Conditions

An important concept in evolutionary game theory is the evolutionary stability of an optimal strategy or equivalently, an equilibrium state. It characterizes the ability of a population to resist small changes or invasions when at equilibrium. Consider a general evolutionary game defined by a payoff matrix  $A$ . Let  $x^* \in S$  be an equilibrium state for the game. Let  $x \in S$  be another arbitrary state. Mix  $x^*$  and  $x \neq x^*$  so that the population changes to a new state,  $\epsilon x + (1 - \epsilon)x^*$ , for some small fraction  $\epsilon > 0$ . Then,  $x^*$  is said to be evolutionarily stable if it remains as a better response to the new “invaded” population state. More accurately, we have the following definition.

**Definition 7.3.4** ((110; 116; 120)). An equilibrium state  $x^* \in S$  for an evolutionary game is evolutionarily stable if there is a small number  $\epsilon' \in (0, 1)$  such that for any  $x \in S$ ,  $x \neq x^*$ ,

$$x^{*T} A(\epsilon x + (1 - \epsilon)x^*) > x^T A(\epsilon x + (1 - \epsilon)x^*), \quad 0 < \epsilon \leq \epsilon'. \quad (7.7)$$

Usually, it is not easy to prove the evolutionary stability of an equilibrium state if just based on its definition. A more straightforward condition is to consider the states  $y$  in a small neighborhood  $U$  of the equilibrium state  $x^*$  and check if no  $y \neq x^*$  prevails  $x^*$  such that  $y^T A y \geq x^{*T} A y$ . It turns out that this condition is necessary and also sufficient:

**Theorem 7.3.5** ((110; 116; 120)). *An equilibrium state  $x^* \in S$  for an evolutionary game is evolutionarily stable if and only if there is a small neighborhood  $U$  of  $x^*$  such that*

$$y^T A y < x^{*T} A y \text{ for all } y \in U \cap S, \quad y \neq x^*. \quad (7.8)$$

Note that a SNgame is an evolutionary game. Therefore, the condition in (7.8) also holds for any of its evolutionarily stable states. For a SNgame,  $x^{*T} A y = y^T A x^*$  since  $A$  is symmetric. Then,  $y^T A y < y^T A x^*$  for all  $y \in U \cap S$ ,  $y \neq x^*$ . It follows that  $y^T A y < x^{*T} A x^*$  for all  $y \in U \cap S$ ,  $y \neq x^*$  since  $y^T A x^* \leq x^{*T} A x^*$  for all  $y \in S$ . This implies that if  $x^*$  is an evolutionarily stable state for a SNgame, it must be a strict local maximizer of the corresponding SPproblem. It turns out that the converse is also true:

**Theorem 7.3.6.** *An equilibrium state  $x^* \in S$  for a SNgame in (7.1) is evolutionarily stable if and only if it is a strict local maximizer of the corresponding SPproblem in (7.4).*

The theorem is actually true for any symmetric evolutionary game (SEgame) and its corresponding generalized knapsack problem (GKproblem), for which a proof can be found in some standard literature (110; 116; 120). Based on this theorem, we can check the stability of an equilibrium state for the SNgame more directly based on the strictness of the state as a local maximizer of the corresponding SPproblem. Though, it is not a simple matter to see if a local maximizer is strict or not. Here, we provide a necessary and sufficient condition for the strictness of a local maximizer from general optimization theory, which we may need to justify some of our results later.

**Theorem 7.3.7** ((105; 111)). *Let  $x^* \in R^n$  be a KKT point for the SPproblem in (7.4). Assume that the active inequalities of the SPproblem at  $x^*$  are all strongly active. Then,  $x^*$  is a strict local maximizer of the SPproblem if and only if  $Z^T AZ$  is negative definite, where  $Z$  is the null space matrix of the Jacobian of the active constraints of the SPproblem at  $x^*$ .*

## 7.4 Cliques as Equilibrium States

As we have mentioned in our introduction section, with a goal of maximizing the social activities, a population would often migrate to sites with the most possible connections among them. Such a group of sites corresponds to a subgraph in the network with a dense set of edges among its nodes. The best one would be a complete subgraph. We call it a social clique. More specifically, we say that a state  $x$  is a social clique if  $x_i > 0$  for all node  $i$  in the clique and  $x_i = 0$  otherwise. In this state, the population is distributed over a clique. However, the cliques are not always optimal choices for a game. Even when they are, they may not necessarily be stable states.

### 7.4.1 Is a Clique an Equilibrium State?

The answer is NOT ALWAYS. See for example the game on graph  $G$  in Fig. 7.3. Notice that the nodes  $\{1, 2, 3\}$  form a clique. If it is an equilibrium state, the values for  $x_1^*$ ,  $x_2^*$ , and  $x_3^*$  should be the same, i.e.,  $x_1^* = x_2^* = x_3^* = 1/3$ , and  $x_i^* = 0$  for all  $i = 4, \dots, 8$ . Then,  $x^{*T} Ax^* = (Ax^*)_i = 2/3$  for all  $i$  such that  $x_i^* > 0$ , i.e.,  $i = 1, 2, 3$ . However,  $x^{*T} Ax^* \geq (Ax^*)_i$  does not hold for all  $i$  such that  $x_i^* = 0$ , i.e.,  $i = 4, \dots, 8$ , since  $x^{*T} Ax^* = 2/3 < (Ax^*)_4 = 1$ . By Theorem 7.3.1,  $x^*$  cannot be an

equilibrium state for the game. In general, let  $C = (V_C, E_C)$  be a clique of graph  $G = (V, E)$ ,  $V_C \subseteq V$  and  $E_C = \{(i, j) : i, j \in V_C\}$ . We then have the following theorem.

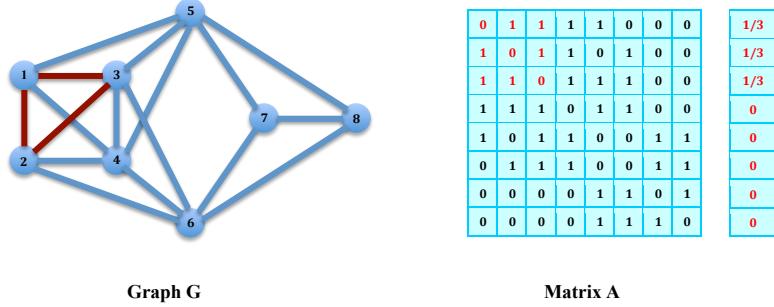


Figure 7.3 Social Cliques: In graph  $G$ , the nodes  $\{1, 2, 3\}$  form a clique. Let  $x^*$  be a strategy for choosing this clique,  $x_1^* = x_2^* = x_3^* = 1/3$  and  $x_i^* = 0$  for all  $i = 4, \dots, 8$ . Then,  $x^{*T}Ax^* = (Ax^*)_i = 2/3$  for all  $i$  such that  $x_i^* > 0$ , i.e.,  $i = 1, 2, 3$ . However,  $x^{*T}Ax^* \geq (Ax^*)_i$  does not hold for all  $i$  such that  $x_i^* = 0$ , i.e.,  $i = 4, \dots, 8$ , since  $x^{*T}Ax^* = 2/3 < (Ax^*)_4 = 1$ . By Theorem 7.3.1,  $x^*$  cannot be an optimal strategy for the SNGame on graph  $G$ .

**Theorem 7.4.1.** *Let  $x^* \in S$  be a strategy on a clique  $C$  of graph  $G$ . Then,  $x^*$  cannot be an optimal strategy for the SNGame in (7.1) if  $C$  is contained in a larger clique of  $G$ .*

*Proof.* If  $x^*$  is optimal for the SNGame, then the population must be distributed evenly over  $C$ , with  $x_i^* = 1/k$  for all  $i \in V_C$  and  $x_i^* = 0$  for all  $i \notin V_C$ , where  $k = |V_C|$ . Since  $A_{i,j} = 1$  for all  $i \neq j \in V_C$  and  $A_{i,i} = 0$  for all  $i \in V_C$ ,  $x^{*T}Ax^* = \sum_{i,j \in V_C} x_i^* A_{i,j} x_j^* = (k-1)/k$ . If  $C$  is contained in a larger clique of  $G$ , there must be a node  $l \in V \setminus V_C$  such that  $(l, j) \in E$  for all  $j \in V_C$ . Then,  $A_{l,j} = 1$  for all  $j \in V_C$ , and  $(Ax^*)_l = \sum_{j \in V_C} A_{l,j} x_j^* = 1 > x^{*T}Ax^*$ . This is a contradiction to the fact that  $(Ax^*)_i \leq x^{*T}Ax^*$  for all  $i \in V$ , as implied by Theorem 7.3.1. Therefore,  $x^*$  cannot be an optimal strategy for the game.  $\square$

The above theorem can also be justified more intuitively as the following: If a clique  $C$  is contained in a larger clique of  $G$ , we can always find a node, say  $l$ , not in  $C$ , and have a strategy  $x$  on this node,  $x = e_l$ , such that  $x^T Ax^* > x^{*T}Ax^*$ , i.e., strategy  $x$  to occupy only node  $l$  prevails strategy  $x^*$  to occupy all the nodes in  $C$ , given the population distributed as  $x^*$  over  $C$ . Then,  $x^*$  is not optimal. Moreover, clique  $C$  and node  $l$  in fact form a larger clique  $H$  of size  $k+1$ , and strategy  $y^*$  on  $H$  has

a better payoff than strategy  $x^*$  on  $C$ , because  $y^{*T}Ay^* = k/(k+1)$  while  $x^{*T}Ax^* = (k-1)/k$ . Therefore, the population would certainly expand from  $C$  to  $H$ , for better social payoff, and so forth.

#### 7.4.2 Is a Maximal Clique an Equilibrium State?

The answer is YES. See for example the game on graph  $G$  again in Fig. 7.4. The nodes  $\{1, 2, 3, 4\}$  form a maximal clique. It is easy to verify that the strategy  $x^*$  on this clique is optimal: If we let  $x_1^* = x_2^* = x_3^* = x_4^* = 1/4$  and  $x_i^* = 0$  for all  $i = 5, \dots, 8$ . Then,  $x^{*T}Ax^* = (Ax^*)_i = 3/4$  for all  $i$  such that  $x_i^* > 0$ , i.e.,  $i = 1, 2, 3, 4$ , and also,  $x^{*T}Ax^* \geq (Ax^*)_i$  for all  $i$  such that  $x_i^* = 0$ , i.e.,  $i = 5, \dots, 8$ . Then,  $x^*$  satisfies all the conditions in Theorem 7.3.1, and it must be an optimal strategy for the game. In general, we have the following theorem.

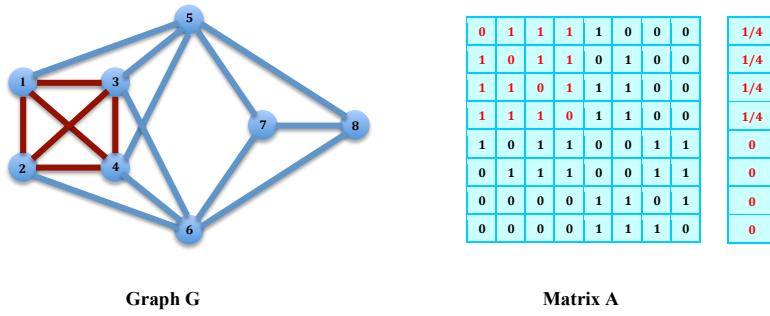


Figure 7.4 Maximal Cliques: The nodes  $\{1, 2, 3, 4\}$  in  $G$  form a maximal clique. Let  $x^*$  be a strategy on this clique,  $x_1^* = x_2^* = x_3^* = x_4^* = 1/4$  and  $x_i^* = 0$  for all  $i = 5, \dots, 8$ . Then,  $x^{*T}Ax^* = (Ax^*)_i = 3/4$  for all  $i$  such that  $x_i^* > 0$ , i.e.,  $i = 1, 2, 3, 4$ , and also,  $x^{*T}Ax^* \geq (Ax^*)_i$  for all  $i$  such that  $x_i^* = 0$ , i.e.,  $i = 5, \dots, 8$ . Then,  $x^*$  satisfies all the conditions in Theorem 7.3.1, and must be an optimal strategy for the SNgame on  $G$ .

**Theorem 7.4.2.** *Let  $C$  be a clique of  $G$ . Let  $x^* \in S$  be a strategy on  $C$ ,  $x_i^* = 1/k$  for all  $i \in V_C$ ,  $x_i^* = 0$  for all  $i \in V \setminus V_C$ , and  $k = |V_C|$ . If  $C$  is a maximal clique, then  $x^*$  is an optimal strategy for the SNgame in (7.1).*

*Proof.* Since  $A_{i,j} = 1$  for all  $i \neq j \in V_C$  and  $A_{i,i} = 0$  for all  $i \in V_C$ , for any  $l \in V_C$ ,  $(Ax^*)_l = \sum_{j \in V_C} A_{l,j} x_j^* = (k-1)/k = x^{*T}Ax^*$ . If  $C$  is a maximal clique of  $G$ , for any  $l \in V \setminus V_C$ , the number of edges from  $l$  to  $C$  is fewer than  $k$ . In other words,  $(l, j) \notin E$  for some  $j \in V_C$ . Then,  $A_{l,j} = 0$  for

some  $j \in V_C$ , and  $(Ax^*)_l = \sum_{j \in V_C} A_{l,j}x_j^* \leq (k-1)/k = x^{*T}Ax^*$ . By Theorem 7.3.1,  $x^*$  must be an optimal strategy for the SNgame in (7.1).  $\square$

Note that an optimal strategy for the SNgame in (7.1) may not always be on a maximal clique. For example, for the game on graph  $G$  in Fig. 7.5, the nodes  $\{1, 2, 3, 4, 5\}$  do not form a clique, but the strategy  $x^*$ ,  $x_1^* = x_3^* = x_4^* = 1/4$ ,  $x_2^* = x_5^* = 1/8$ , and  $x_i^* = 0$  for  $i = 6, 7, 8$ , is in fact optimal: It is easy to verify that  $x^{*T}Ax^* = (Ax^*)_i = 3/4$  for all  $i$  such that  $x_i^* > 0$ , i.e.,  $i = 1, 2, 3, 4, 5$ , and also,  $x^{*T}Ax^* \geq (Ax^*)_i$  for all  $i$  such that  $x_i^* = 0$ , i.e.,  $i = 6, 7, 8$ . Then,  $x^*$  satisfies all the conditions in Theorem 7.3.1, and must be an optimal strategy for the game. We discuss non-clique optimal strategies in greater detail later.

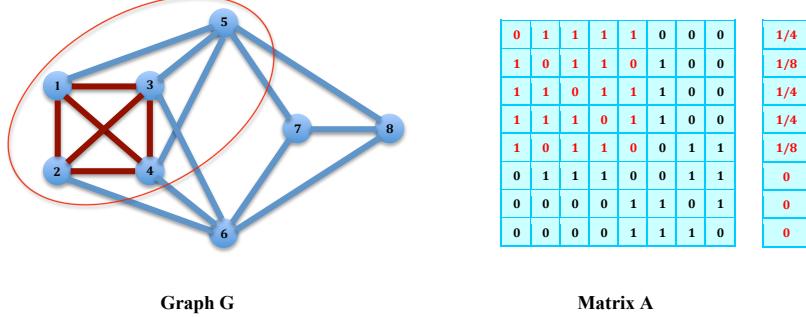


Figure 7.5 Non-Clique Strategies: The nodes  $\{1, 2, 3, 4, 5\}$  in  $G$  do not form a clique, but the strategy  $x^*$ ,  $x_1^* = x_3^* = x_4^* = 1/4$ ,  $x_2^* = x_5^* = 1/8$ , and  $x_i^* = 0$  for  $i = 6, 7, 8$ , is in fact optimal: It is easy to verify that  $x^{*T}Ax^* = (Ax^*)_i = 3/4$  for all  $i$  such that  $x_i^* > 0$ , i.e.,  $i = 1, 2, 3, 4, 5$ , and also,  $x^{*T}Ax^* \geq (Ax^*)_i$  for all  $i$  such that  $x_i^* = 0$ , i.e.,  $i = 6, 7, 8$ . Then,  $x^*$  satisfies all the conditions in Theorem 7.3.1, and must be an optimal strategy for the SNgame on  $G$ .

### 7.4.3 Is a Maximal Clique a Local Maximizer?

A maximal clique is an equilibrium state for the SNgame in (7.1), but it may or may not be stable, which is important to know in practice. By Theorem 7.3.6, an equilibrium state for the SNgame in (7.1) is evolutionarily stable if and only if it is a strict local maximizer of the SPproblem in (7.4). Unfortunately, in some circumstances, a maximal clique may not even be a local maximizer of the SPproblem in (7.4). For example, for the game on graph  $G$  in Fig. 7.6, the nodes  $\{1, 2, 3, 4\}$  form a maximal clique. It is an equilibrium state for the game, but not a local maximizer of the corresponding

SPproblem: Let  $x^*$  represent the equilibrium state on this clique,  $x_i^* = 1/4$  for all  $i = 1, \dots, 4$  and  $x_i^* = 0$  for all  $i = 5, \dots, 8$ . Construct a new state  $x = x^* + p$ , where  $p = (-2\epsilon, 0, 0, 0, \epsilon, \epsilon, 0, 0)^T$  for a small  $\epsilon > 0$ . Then, it is easy to verify that  $x_i \geq 0$  for all  $i = 1, \dots, 8$ ,  $\sum_i x_i = 1$ , and

$$x^T Ax = x^{*T} Ax^* + p^T Ap = x^{*T} Ax^* + 2\epsilon^2 > x^{*T} Ax^*,$$

for all small  $\epsilon > 0$ . As  $\epsilon$  goes to zero,  $x$  is arbitrarily close to  $x^*$ , yet  $x^T Ax > x^{*T} Ax^*$ . Therefore,  $x^*$  cannot be a local maximizer for the SPproblem in (7.4). Notice in this example that in  $G$ , there is a bigger clique,  $\{2, 3, 4, 5, 6\}$ , attaching to 3 nodes of the clique  $\{1, 2, 3, 4\}$ , which is in fact a critical factor for why the latter cannot be a local maximizer for the SPproblem in (7.4). In general we have the following theorems.

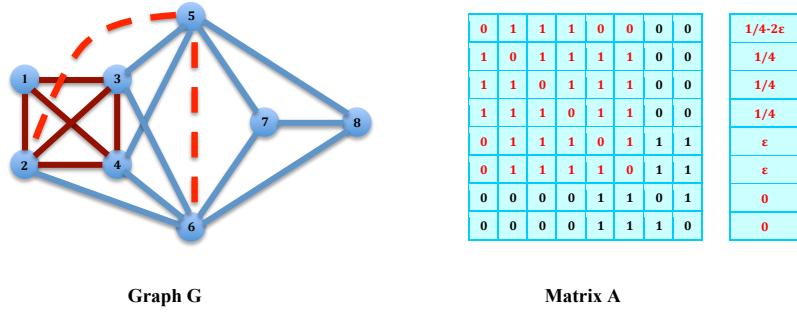


Figure 7.6 Maximal Cliques vs. Local Maximizers: The nodes  $\{1, 2, 3, 4\}$  in  $G$  form a maximal clique. It is an equilibrium state for the SNgame on  $G$ , but not a local maximizer of the corresponding SPproblem: Let  $x^*$  be the equilibrium state on this clique,  $x_i^* = 1/4$  for all  $i = 1, \dots, 4$  and  $x_i^* = 0$  for all  $i = 5, \dots, 8$ . Construct a new state  $x = x^* + p$ , where  $p = (-2\epsilon, 0, 0, 0, \epsilon, \epsilon, 0, 0)^T$  for a small  $\epsilon > 0$ . Then, it is easy to verify that  $x_i \geq 0$  for all  $i = 1, \dots, 8$ ,  $\sum_i x_i = 1$ , and  $x^T Ax = x^{*T} Ax^* + 2\epsilon^2 > x^{*T} Ax^*$  for all small  $\epsilon > 0$ . As  $\epsilon$  goes to zero,  $x$  is arbitrarily close to  $x^*$ , yet  $x^T Ax > x^{*T} Ax^*$ . Therefore,  $x^*$  cannot be a local maximizer for the corresponding SPproblem.

**Theorem 7.4.3.** *Let  $C$  be a maximal clique of graph  $G$ . Let  $x^* \in S$  be a strategy on  $C$  for the SNgame in (7.1),  $x_i^* = 1/k$  for all  $i \in V_C$ ,  $x_i^* = 0$  for all  $i \in V \setminus V_C$ , and  $k = |V_C|$ . Then,  $x^*$  is a local maximizer for the SPproblem in (7.4), only if there is no larger clique attaching to  $k - 1$  nodes of  $C$ .*

*Proof.* Assume that  $x^*$  is a local maximizer for the SPproblem in (7.4), but there is a larger clique  $H$  attaching to  $k - 1$  nodes of  $C$ . Assume without loss of generality that  $H$  is of size  $k + 1$ . Then,

there must be two nodes  $i, j \in V_H \setminus V_C$  each connected with the same set of  $k - 1$  nodes of  $C$ , and both are also connected to each other, i.e.,  $(i, j) \in E_H$ . Let  $l$  be the node in  $C$  not connected with node  $i$  and  $j$ . Define a vector  $p \in R^n$ , with  $p_l = -2\epsilon$ ,  $p_i = p_j = \epsilon$ , and  $p_r = 0$  for all  $r \neq i, j, l$ , where  $0 \leq \epsilon \leq 1/2k$ . Then,  $x = x^* + p$  satisfies all the constraints for the SPproblem in (7.4). Note that  $(Ax^*)_l = (Ax^*)_i = (Ax^*)_j = (k - 1)/k$ . Therefore,  $p^T Ax^* = 0$ . It follows that  $x^T Ax = x^{*T} Ax^* + p^T Ap = x^{*T} Ax^* + 2\epsilon^2 > x^{*T} Ax^*$  for all  $\epsilon > 0$ . Then,  $x^*$  cannot be a local maximizer of the SPproblem in (7.4).  $\square$

In fact, the above condition for not having a larger attached clique is also sufficient for the strategy on a maximal clique for the SNgame to be a local maximizer of the corresponding SPproblem. However, the proof is more mathematically involved. We provide a justification in the following.

**Theorem 7.4.4.** *Let  $C$  be a maximal clique of graph  $G$ . Let  $x^* \in S$  be a strategy on  $C$  for the SNgame in (7.1),  $x_i^* = 1/k$  for all  $i \in V_C$ ,  $x_i^* = 0$  for all  $i \in V \setminus V_C$ , and  $k = |V_C|$ . Then,  $x^*$  is a local maximizer for the SPproblem in (7.4), if there is not larger clique attaching to  $k - 1$  nodes of  $C$ .*

*Proof.* Without loss of generality, let  $V_C = \{1, \dots, k\}$ . Then,  $x_i^* = 1/k$  for all  $i = 1, \dots, k$  and  $x_i^* = 0$  for all  $i > k$ . We show that there is a small neighborhood  $U$  of  $x^*$  such that  $x^T Ax \leq x^{*T} Ax^*$  for all  $x$  feasible in  $U$ .

Let  $U = \{x \in R^n : \|x - x^*\| < r\}$  for some  $r > 0$ . If  $x \in U$ , then  $x = x^* + \epsilon$  for some  $\epsilon \in R^n$ ,  $\|\epsilon\| < r$ . Let  $S = \{x \in R^n : \sum_i x_i = 1, x_i \geq 0\}$ . If  $x \in S$ , then  $x = x^* + \epsilon$  for some  $\epsilon \in R^n$ ,  $\sum_i \epsilon_i = 0$ ,  $\epsilon_i \geq -1/k$  for all  $i = 1, \dots, k$ , and  $\epsilon_i \geq 0$  for all  $i > k$ .

Let  $x \in U \cap S$ . Define vector  $x^+ \in R^k$ ,  $\epsilon^+ \in R^k$ , and  $\epsilon^0 \in R^{n-k}$ , with  $x_i^+ = 1/k$ ,  $i = 1, \dots, k$ . We can then write  $x^* = (x^+, 0)^T$ ,  $\epsilon = (\epsilon^+, \epsilon^0)^T$ , and  $x = (x^+ + \epsilon^+, \epsilon^0)^T$ , with  $\epsilon_i^+ \geq -1/k$  for all  $i = 1, \dots, k$ ,  $\epsilon_i^0 \geq 0$  for all  $i = 1, \dots, n - k$ ,  $\sum_i \epsilon_i^+ + \sum_i \epsilon_i^0 = 0$ , and  $\|\epsilon\| < r$  for some  $r > 0$ .

Let  $A$  be decomposed into four submatrices,  $\{A^{(i,j)} : i, j = 1, 2\}$  with  $A^{(1,1)} = A_{1:k, 1:k}$ ,  $A^{(1,2)} = A_{1:k, k+1:n}$ ,  $A^{(2,1)} = A_{k+1:n, 1:k}$ , and  $A^{(2,2)} = A_{k+1:n, k+1:n}$ . Then,

$$\begin{aligned} x^T Ax &= x^{+T} A^{(1,1)} x^+ + 2x^{+T} A^{(1,1)} \epsilon^+ + \epsilon^{+T} A^{(1,1)} \epsilon^+ \\ &\quad + 2x^{+T} A^{(1,2)} \epsilon^0 + 2\epsilon^{+T} A^{(1,2)} \epsilon^0 + \epsilon^{0T} A^{(2,2)} \epsilon^0. \end{aligned}$$

Since  $C$  is a clique of  $G$ , the elements in  $A^{(1,1)}$  are all equal to 1 except for those along the diagonal, and therefore,  $x^{*T}Ax^* = x^{+T}A^{(1,1)}x^+ = (k-1)/k$ . So, in order to prove  $x^TAx \leq x^{*T}Ax^*$ , all we need to do is to show

$$\begin{aligned} & 2x^{+T}A^{(1,1)}\epsilon^+ + \epsilon^{+T}A^{(1,1)}\epsilon^+ + \\ & 2x^{+T}A^{(1,2)}\epsilon^0 + 2\epsilon^{+T}A^{(1,2)}\epsilon^0 + \epsilon^{0T}A^{(2,2)}\epsilon^0 \leq 0. \end{aligned} \quad (7.9)$$

or equivalently,

$$\begin{aligned} & \epsilon^{+T}A^{(1,1)}\epsilon^+ + 2\epsilon^{+T}A^{(1,2)}\epsilon^0 + \epsilon^{0T}A^{(2,2)}\epsilon^0 \\ & \leq -2x^{+T}A^{(1,1)}\epsilon^+ - 2x^{+T}A^{(1,2)}\epsilon^0. \end{aligned} \quad (7.10)$$

Let  $\sum_i \epsilon_i^0 = s_0$ . Then  $s_0 \geq 0$ , and  $\sum_i \epsilon_i^+ = -s_0$ . It follows that  $-2x^{+T}A^{(1,1)}\epsilon^+ = 2(k-1)s_0/k$ . However, since  $C$  is a maximal clique, there are at most  $k-1$  elements equal to 1 in each of the columns of  $A^{(1,2)}$ . Therefore,  $2x^{+T}A^{(1,2)}\epsilon^0 \leq 2(k-1)s_0/k$ . We now consider two cases of this inequality separately, one for strictly less than ( $<$ ) and another for exactly equal to ( $=$ ):

(i) First assume  $2x^{+T}A^{(1,2)}\epsilon^0 < 2(k-1)s_0/k$ . Then, the right-hand side of inequality (7.10) is greater than zero, and

$$-2x^{+T}A^{(1,1)}\epsilon^+ - 2x^{+T}A^{(1,2)}\epsilon^0 = Ls_0 = O(s_0) = O(\|\epsilon\|),$$

for some constant  $L > 0$ . On the other hand, the left-hand side of inequality (7.10) is basically an expanded form of  $\epsilon^T A \epsilon$ . Therefore,

$$\epsilon^{+T}A^{(1,1)}\epsilon^+ + 2\epsilon^{+T}A^{(1,2)}\epsilon^0 + \epsilon^{0T}A^{(2,2)}\epsilon^0 = \lambda\|\epsilon\|^2 = O(\|\epsilon\|^2),$$

where  $\lambda$  is a number between the largest and smallest eigenvalues of  $A$ . It follows that there must be a number  $r > 0$  such that when  $\|\epsilon\| \leq r$ , the left-hand side of (7.10) will always be smaller than the right-hand side of (7.10).

(ii) Now assume  $2x^{+T}A^{(1,2)}\epsilon^0 = 2(k-1)s_0/k$ . Inequality (7.10) then becomes

$$\epsilon^{+T}A^{(1,1)}\epsilon^+ + 2\epsilon^{+T}A^{(1,2)}\epsilon^0 + \epsilon^{0T}A^{(2,2)}\epsilon^0 \leq 0. \quad (7.11)$$

Let  $\bar{A}^{(1,1)} = ee^T - A^{(1,1)}$ , where  $e$  is a vector of all 1's in  $R^k$ ,  $\bar{A}^{(2,2)} = ff^T - A^{(2,2)}$ , where  $f$  is a vector of all 1's in  $R^{n-k}$ , and  $\bar{A}^{(1,2)} = ef^T - A^{(1,2)}$ . Then, inequality (7.11) transforms to

$$-\epsilon^{+T}\bar{A}^{(1,1)}\epsilon^+ - 2\epsilon^{+T}\bar{A}^{(1,2)}\epsilon^0 - \epsilon^{0T}\bar{A}^{(2,2)}\epsilon^0 \leq 0. \quad (7.12)$$

Since the elements of  $A^{(1,1)}$  are all 1's except for those along the diagonal,  $\bar{A}^{(1,1)}$  is an identity matrix, and

$$-\epsilon^{+T} \bar{A}^{(1,1)} \epsilon^+ = -\sum_i \epsilon_i^{+2}. \quad (7.13)$$

Also, since  $2x^{+T} A^{(1,2)} \epsilon^0 = 2(k-1)s_0/k$ , there must be exactly  $k-1$  elements equal to 1 in each of the columns of  $A^{(1,2)}$  and hence one element equal to 1 in each of the columns of  $\bar{A}^{(1,2)}$ . Then,

$$\begin{aligned} -2\epsilon^{+T} \bar{A}^{(1,2)} \epsilon^0 &= -2 \sum_i \epsilon_i^+ (\sum_{j \in D_i} \epsilon_j^0) \leq \sum_i \epsilon_i^{+2} + \sum_i (\sum_{j \in D_i} \epsilon_j^0)^2 \\ &= \sum_i \epsilon_i^{+2} + \sum_j \epsilon_j^{02} + \sum_l \sum_{i \neq j \in D_l} \epsilon_i^0 \epsilon_j^0, \end{aligned} \quad (7.14)$$

where  $D_l$  is the set of nodes in  $V \setminus V_C$  connecting to the same  $k-1$  nodes of  $C$  except for node  $l \in V_C$ .

Now, since there is no larger clique attaching to  $k-1$  nodes of  $C$ , any two nodes  $i \neq j \in V \setminus V_C$  connecting to the same  $k-1$  nodes in  $C$  do not have a link between them and therefore,  $A_{i,j}^{(2,2)} = A_{j,i}^{(2,2)} = 0$ , and  $\bar{A}_{i,j}^{(2,2)} = \bar{A}_{j,i}^{(2,2)} = 1$  for all  $i \neq j \in D_l$ ,  $l \in V_C$ . Given the fact that all diagonal elements of  $\bar{A}^{(2,2)}$  are 1's,

$$\epsilon^{0T} \bar{A}^{(2,2)} \epsilon^0 \geq \sum_j \epsilon_j^{02} + \sum_l \sum_{i \neq j \in D_l} \epsilon_i^0 \epsilon_j^0. \quad (7.15)$$

By combining (7.13), (7.14), and (7.15), the inequality (7.12) then follows.  $\square$

As a KKT point for the SPproblem in (7.4), a maximal clique strategy  $x^*$  for the SNgame in (7.1) may be a local minimizer, saddle point, or local maximizer (107; 114). Only when it is a strict local maximizer, it is evolutionarily stable. When it is just a local maximizer, it is weakly evolutionarily stable in the sense that it will not be a worse response than any other strategy  $x$  in an “invaded” population  $(\epsilon x + (1-\epsilon)x^*)$  (110; 120). Formally, we have the following definition for weak evolutionary stability:

**Definition 7.4.5** ((110; 116; 120)). An equilibrium state  $x^* \in S$  for an evolutionary game is weakly evolutionarily stable if there is a small number  $\epsilon' \in (0, 1)$  such that for any  $x \in S$ ,  $x \neq x^*$ ,

$$x^{*T} A(\epsilon x + (1-\epsilon)x^*) \geq x^T A(\epsilon x + (1-\epsilon)x^*), \quad 0 < \epsilon \leq \epsilon'. \quad (7.16)$$

Now, the next question would be under what condition a maximal clique strategy can be a strict local maximizer of the SPproblem, and we can then claim its evolutionary stability. We answer this question in the following subsection.

#### 7.4.4 Is a Maximal Clique a Strict Local Maximizer?

A maximal clique strategy for the SNGame in (7.1) is a local maximizer of the SPproblem in (7.4) unless there is a larger clique attached to it. It is also a strict local maximizer except in some special situations. For example, in graph  $G$  in Fig. 7.7, the nodes  $\{1, 2, 3, 4\}$  form a maximal clique. Since it is not attached with any larger clique with any 3 of its nodes, the strategy  $x^*$  on this clique for the SNGame on  $G$ ,  $x_i^* = 1/4$  for all  $i = 1, \dots, 4$  and  $x_i^* = 0$  for all  $i = 5, \dots, 8$ , is a local maximizer of the corresponding SPproblem, and  $x^{*T}Ax^* \geq x^T Ax$  for any  $x \in S$  in a small neighborhood  $U$  of  $x^*$ . However, if we choose  $x \neq x^*$  such that  $x_1 = x_3 = x_4 = 1/4$ ,  $x_2 = 1/4 - \epsilon$ ,  $x_5 = \epsilon$ , and  $x_i = 0$  for  $i = 6, 7, 8$ , we see for any  $U$  that  $x \in S \cap U$  for sufficiently small  $\epsilon > 0$ , and  $x^T Ax = x^{*T}Ax^* = 3/4$ . Therefore,  $x^*$  cannot be a strict local maximizer for the SPproblem. Notice that in graph  $G$ , node 5 is connected with 3 nodes, node 1, 3, 4, of clique  $\{1, 2, 3, 4\}$ , which makes it possible to construct the strategy  $x \in S \cap U$  on node 1, 2, 3, 4, 5 so that  $x^T Ax = x^{*T}Ax^*$ . In general, we have the following necessary and sufficient conditions for a maximal clique strategy to be a strict local maximizer of the SPproblem.

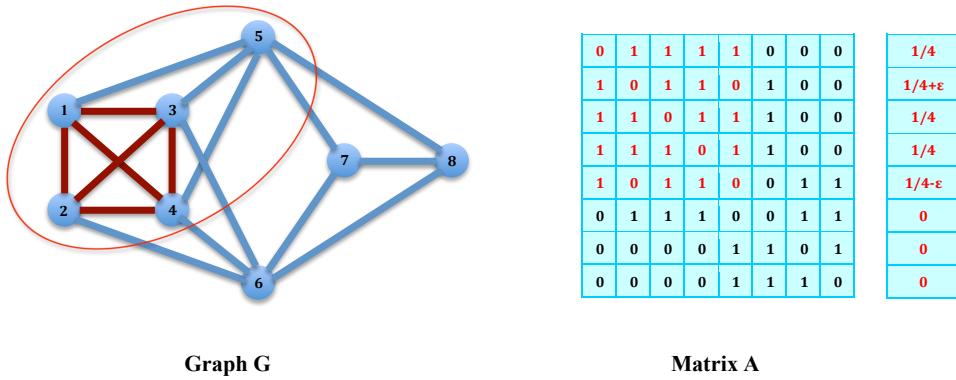


Figure 7.7 Strict Local Maximizer: The nodes  $\{1, 2, 3, 4\}$  in  $G$  form a maximal clique. The strategy  $x^*$  on this clique for the SNGame on  $G$ ,  $x_i^* = 1/4$  for all  $i = 1, \dots, 4$  and  $x_i^* = 0$  for all  $i = 5, \dots, 8$ , is a local maximizer of the corresponding SPproblem, and  $x^{*T}Ax^* \geq x^T Ax$  for any  $x \in S$  in a small neighborhood  $U$  of  $x^*$ . However, if we choose  $x \neq x^*$  such that  $x_1 = x_3 = x_4 = 1/4$ ,  $x_2 = 1/4 - \epsilon$ ,  $x_5 = \epsilon$ , and  $x_i = 0$  for  $i = 6, 7, 8$ , we see for any  $U$  that  $x \in S \cap U$  for sufficiently small  $\epsilon > 0$ , and  $x^T Ax = x^{*T}Ax^* = 3/4$ . Therefore,  $x^*$  is not a strict local maximizer for the SPproblem.

**Theorem 7.4.6.** *Let  $C$  be a maximal clique of graph  $G$ . Let  $x^* \in S$  be a strategy on  $C$  for the SNgame in (7.1),  $x_i^* = 1/k$  for all  $i \in V_C$ ,  $x_i^* = 0$  for all  $i \in V \setminus V_C$ , and  $k = |V_C|$ . Then,  $x^*$  is a strict local maximizer for the SPproblem in (7.4), only if there is no node in  $V \setminus V_C$  connected to  $k - 1$  nodes of  $C$ .*

*Proof.* Assume that  $x^*$  is a strict local maximizer for the SPproblem in (7.4), and there is a node  $l \in V \setminus V_C$  connected to  $k - 1$  nodes of  $C$ , i.e.,  $(l, j) \in E \setminus E_C$  for  $k - 1$  node  $j \in V_C$ . Then,  $(Ax^*)_l = (k - 1)/k$ . Let  $l' \in V_C$  be the node that node  $l$  is not connected to. Since  $l' \in V_C$ ,  $(Ax^*)_{l'} = (k - 1)/k$ . Now, construct a strategy  $x = x^* + p$ , with  $p_l = \epsilon$ ,  $p_{l'} = -\epsilon$ , and  $p_i = 0$  for all  $i \neq l, l'$ . Then,  $x \in S \cap U$  for sufficiently small  $\epsilon > 0$ , while

$$x^T Ax = x^{*T} Ax^* + 2p^T Ax^* + p^T Ap = x^{*T} Ax^*,$$

since  $p^T Ax^* = \epsilon(Ax^*)_l - \epsilon(Ax^*)_{l'} = 0$ , and  $p^T Ap = (A_{l,l} - 2A_{l,l'} + A_{l',l'})\epsilon^2 = 0$ . This is a contradiction to the assumption that  $x^*$  is a strict local maximizer for the SPproblem. So there cannot be any node in  $V \setminus V_C$  connected to  $k - 1$  nodes of  $C$ .  $\square$

**Theorem 7.4.7.** *Let  $C$  be a maximal clique of graph  $G$ . Let  $x^* \in S$  be a strategy on  $C$  for the SNgame in (7.1),  $x_i^* = 1/k$  for all  $i \in V_C$ ,  $x_i^* = 0$  for all  $i \in V \setminus V_C$ , and  $k = |V_C|$ . Then,  $x^*$  is a strict local maximizer for the SPproblem in (7.4), if there is no node in  $V \setminus V_C$  connected to  $k - 1$  nodes of  $C$ .*

*Proof.* Without loss of generality, let's assume that  $V_C = \{1, \dots, k\}$ . Then,  $x_i^* = 1/k$  for all  $i = 1, \dots, k$  and  $x_i^* = 0$  for all  $i > k$ . Since there is no node in  $V \setminus V_C$  connected to  $k - 1$  nodes of  $C$ , there is no larger clique attaching to any of  $k - 1$  nodes of  $C$ , either. It follows from Theorem 7.4.4 that  $x^*$  is a local maximizer of the SPproblem in (7.4). Since there is no node in  $V \setminus V_C$  connected to  $k - 1$  nodes of  $C$ , the active constraints of the SPproblem at  $x^*$  are all strongly active. Therefore, by Theorem 7.3.7, in order to prove that  $x^*$  is a strict local maximizer of the SPproblem, all we need to show is that  $Z^T AZ$  is negative definite, where  $Z$  is a null space matrix of the Jacobian matrix of the active constraints of the SPproblem at  $x^*$ .

The active constraints of the SPproblem at  $x^*$  include all  $x_i \geq 0$  for  $i > k$  and  $\sum_i x_i = 0$ . Then, the null space matrix  $Z$  for the Jacobian of this set of constraints is an  $n \times (k - 1)$  matrix and can be defined as follows:  $Z_{1,j} = -1$  and  $Z_{j+1,j} = 1$  for all  $j = 1, \dots, (k - 1)$ , and all other elements  $Z_{i,j} = 0$  (111). Let  $\bar{A}$  be the first  $k$  rows and  $k$  columns of  $A$ , and  $\bar{Z}$  the first  $k$  rows of  $Z$ . Then,  $Z^T AZ = \bar{Z}^T \bar{A} \bar{Z}$ .

Since  $C$  is a clique, the elements of  $\bar{A}$  are all 1's but 0's along the diagonal. Therefore,  $\bar{A} = ee^T - I$ , where  $e \in R^k$  is a vector of all 1's and  $I$  a  $k \times k$  identity matrix. Then,

$$Z^T AZ = \bar{Z}^T \bar{A} \bar{Z} = \bar{Z}^T (ee^T - I) \bar{Z} = -\bar{Z}^T \bar{Z}$$

is negative definite, proving that  $x^*$  is a strict local maximizer of the SPproblem in (7.4).  $\square$

We have now learned that a strategy on a maximal clique  $C$  of size  $k$  is optimal for a SNgame, when the population is distributed evenly on the clique, and the network reaches equilibrium. The equilibrium is at least weakly evolutionarily stable unless there is a larger clique attached to  $k - 1$  nodes of  $C$ . In the latter case, the population would in some sense be able to migrate easily to the neighboring larger clique, to achieve a higher social payoff. It is evolutionarily stable unless there is a node not in  $C$  but connected to  $k - 1$  nodes of  $C$  or equivalently, there is a clique of size  $k$  attached to  $k - 1$  nodes of  $C$ . In the latter case, again, the population would in some sense be able to swing easily to the neighboring clique without losing any social payoff.

#### 7.4.5 Is a Maximum Clique an Optimal and Stable Strategy?

Given a graph  $G$ , the maximum clique is a maximal clique of the largest size in  $G$ . Is a strategy on such a clique an optimal strategy for a SNgame on  $G$ ? If yes, is it evolutionarily stable? The answer for the first question is easy: Since the maximum clique is a maximal clique, the strategy is certainly optimal. In addition, since there is no larger cliques in  $G$  than the maximum clique, there cannot be any larger clique attaching to it, and by Theorem 7.4.3 and 7.4.4, the strategy must be a local maximizer of the corresponding SPproblem and be at least weakly evolutionarily stable. However, there could be a node which is not in the clique, but connects to  $k - 1$  nodes of the clique, for a maximum clique of size  $k$ . Therefore, by Theorem 7.4.6 and 7.4.7, the strategy can be a strict local maximizer and hence be evolutionarily stable unless there is a node not in the clique but connected to  $k - 1$  nodes of the clique. For example, in graph  $G$  in Fig. 7.7, the nodes  $\{1, 2, 3, 4\}$  form a maximum clique. The strategy on this clique is optimal, but it is not a strict local maximizer of the corresponding SPproblem because node 5 connects to 3 nodes of the clique, node 2, 3, 4, and therefore, it is not evolutionarily stable.

Let  $C$  be a maximum clique of size  $k$  of graph  $G$ . Then, the strategy  $x^*$  on  $C$  for the SNgame in (7.1),  $x_i^* = 1/k$  for all  $i \in V_C$  and  $x_i^* = 0$  for all  $i \in V \setminus V_C$ , is a local maximizer of the SPproblem

in (7.4), with the maximum equal to  $x^{*T}Ax^*/2 = (k-1)/2k$ . However, Motzkin and Straus 1965 showed that the global maximum of a SPproblem is always equal to  $(k-1)/2k$ , where  $k$  is the size of the maximum clique of  $G$  (113). Therefore,  $x^*$  on  $C$  must be a global maximizer of the SPproblem in (7.4). In general, we have the following two theorems:

**Theorem 7.4.8.** *Let  $H$  be a subgraph of  $G$  and  $x^*$  a strategy on  $H$ ,  $x^* > 0$  for all  $i \in V_H$  and  $x^* = 0$  for all  $i \in V \setminus V_H$ . If  $x^*$  is a global maximizer of the SPproblem in (7.4), then the global maximum  $x^{*T}Ax^*/2$  of the problem must equal to  $(k-1)/2k$ , where  $k$  is the size of the maximum clique  $C$  contained in  $H$ .*

**Theorem 7.4.9.** *The strategy  $x^*$  on a maximum clique  $C$  of  $G$  for the SNgame in (7.1) is a global maximizer of the SPproblem in (7.4) with the global maximum equal to  $(k-1)/2k$ , where  $k$  is the size of  $C$ .*

Note that a global maximizer may not necessarily be on a maximum clique, as given in the above theorem. In particular, a global maximizer can be in a form with more than  $k$  nonzero components including those corresponding to the nodes of a maximum clique. However, once a global maximizer is obtained, we can always apply a simple procedure, as described below, to recover one that corresponds to a maximum clique.

Let  $x^* \in R^n$  be a global maximizer of the SPproblem in (7.4) with the maximum equal to  $(k-1)/2k$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , where  $H$  is a subgraph of  $G$ , and  $l = |H| \geq k$ . Based on Motzkin and Straus 1965,  $H$  must contain a maximum clique of size  $k$  (113). If  $l > k$ , then  $H$  must be incomplete, and there must be  $i, j \in V_H$ ,  $(i, j) \notin E_H$ .

Define  $p \in R^n$ , with  $p_i = c$ ,  $p_j = -c$ , and  $p_r = 0$  for all  $r \neq i, j$ , where  $-x_i^* \leq c \leq x_j^*$ . Then  $y^* = x^* + p$  remains feasible for all the constraints of the SPproblem in (7.4), and

$$y^{*T}Ay^* = (x^* + p)^T A(x^* + p) = x^{*T}Ax^* + 2x^{*T}Ap + p^TAp.$$

Since  $x^*$  is a global maximizer of the SPproblem in (7.4), it must satisfy the KKT conditions of the SPproblem in (7.4). In particular, there must be a number  $\lambda^*$  such that  $(Ax^*)_r = \lambda^*$  for all  $r \in V_H$ . It follows that  $2x^{*T}Ap = 2\lambda^*(c - c) = 0$ . Since  $(i, j) \notin E_H$ ,  $A_{i,j} = A_{j,i} = 0$ , and  $p^TAp = 0$ . It follows that  $y^{*T}Ay^* = x^{*T}Ax^*$ , and  $y^*$  remains to be a global maximizer of the SPproblem in (7.4). If we set

$c = x_j^*$  (or  $c = -x_i^*$ ), we then obtain  $y^*$  with fewer nonzero components, corresponding to a subgraph of  $H$  with node  $j \in V_H$  (or  $i \in V_H$ ) and its connected edges eliminated.

The above procedure can be repeated for the new maximizer  $y^*$  and reduced subgraph  $H$  until a complete subgraph of size  $k$  is reached. The whole process would require only  $l - k$  steps, for there are only  $l - k$  nodes to be eliminated from a starting subgraph  $H$ . An example is given in Fig. 7.8 to demonstrate how the procedure works. In the first graph on the top, the nodes  $\{1, 2, 3, 4, 5, 6\}$  form a subgraph  $H$ , and  $x^* \in S$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , is a global maximizer for the SPproblem. However,  $H$  is not a maximum clique. Since node 2 and 6 are not connected, we therefore add 1/8 to  $x_2$  but subtract it from  $x_6$ . We then obtain a reduced subgraph  $H$ , with  $V_H = \{1, 2, 3, 4, 5\}$ ,  $E_H = \{(i, j) \in E : i, j \in V_H\}$ , as shown in the second graph to the top. The solution  $x^* \in S$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , remains to be a global maximizer for the SPproblem. Next, since node 1 and 5 are still not connected, we then add 1/8 to  $x_1$  but subtract it from  $x_5$  to obtain a further reduced subgraph  $H$ , with  $V_H = \{1, 2, 3, 4\}$ ,  $E_H = \{(i, j) \in E : i, j \in V_H\}$ , as shown in the graph in the bottom. The solution  $x^* \in S$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , again remains to be a global maximizer for the SPproblem, but this time,  $H$  is a maximum clique of  $G$ .

Based on Theorem 7.4.8 and 7.4.9 and the above graph reduction procedure, it is clear that the maximum clique of a given graph  $G$  can be found by solving a SPproblem to its global maximum and then followed by performing a graph reduction process. The latter requires only linear time, therefore, solving a SPproblem to its global maximum is at least as hard as the problem of finding a maximum clique for a graph  $G$ . We know that the latter problem is NP-hard (104; 108). It follows that the problem to find a global maximizer for the SPproblem in (7.4) must be NP-hard. This fact has been aware to some extent in the field of evolutionary game theory, but not been formally justified. We state it as a theorem in the following. The proof follows straightforwardly from our discussions.

**Theorem 7.4.10.** *The problem to find a global maximizer for the SPproblem in (7.4) and equivalently, the problem to find an equilibrium state of maximum payoff for the SNgame in (7.1) is NP-hard.*

Note that the graph reduction procedure described above can also be reversed, i.e., starting from a global maximizer of the SPproblem on a maximum clique, we can extend it to a global maximizer

on a general subgraph that contains the clique. Key to this process is to make sure that in every step, an extended subgraph can be found such that every pairs of unconnected nodes in the subgraph are connected with the same set of nodes in the subgraph (as justified in Theorem 7.4.11). For example, in Fig. 7.8, the subgraph formed by nodes  $\{1, 2, 3, 4, 5\}$  in the second graph from the bottom has two nodes separated, node 1 and 5, and both are connected with the same set of nodes in the subgraph,  $\{2, 3, 4\}$ . Let  $x^*$  be the global maximizer for the graph at the bottom of the figure. Then, by adding a small number  $c = 1/8$  to  $x_5^*$  and subtract it from  $x_1^*$ , we obtain a global maximizer on the subgraph formed by the nodes  $\{1, 2, 3, 4, 5\}$ . Likewise, we can further extend this subgraph to  $\{1, 2, 3, 4, 5, 6\}$  to obtain another global maximizer for the problem as shown in the graph on the top of the figure.

**Theorem 7.4.11.** *Let  $x^* \in R^n$  be a global maximizer for the SPproblem in (7.4),  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , where  $H$  is a subgraph of  $G$ , with node  $l, l' \in V_H$  not connected. Then, node  $l$  and  $l'$  must connect to the same set of nodes in  $V_H$ .*

*Proof.* Since  $x^*$  is a global maximizer for the SPproblem in (7.4),  $x^{*T} Ax^* = (k - 1)/2k$ , where  $k$  is the size of the maximum clique contained in  $H$ . Define  $p \in R^n$ ,  $p_l = c$ ,  $p_{l'} = -c$ , and  $p_r = 0$  for all  $r \neq l, l'$ , with  $-x_l \leq c \leq x_{l'}$ . Then,  $x = x^* + p$  is also a global maximizer, and  $x^T Ax = x^{*T} Ax^* = (k - 1)/2k$ . Then, by Theorem 7.3.2,  $(Ax)_i = x^T Ax$  for all  $i \in V_H$ ,  $i \neq l, l'$ . However,

$$(Ax)_i = \sum_{j \in V_H} A_{i,j} x_j = \sum_{j \in V_H} A_{i,j} x_j^* + A_{i,l} c - A_{i,l'} c = x^{*T} Ax^* + A_{i,l} c - A_{i,l'} c.$$

It follows that  $A_{i,l} c - A_{i,l'} c = 0$  and  $A_{i,l} = A_{i,l'}$  for all  $i \in V_H$ ,  $i \neq l, l'$ , i.e., node  $l$  and  $l'$  are connected to the same set of nodes in  $H$ .  $\square$

#### 7.4.6 Non-Clique Equilibrium States

There could be many non-clique equilibrium states for a given SNgame. Some of them can be derived directly from the equilibrium states on maximal cliques as shown in the example in Fig. 7.5. In general, they are not as stable as the states on cliques. For example, in Fig. 7.8, the optimal strategies shown in the top two graphs are non-clique equilibrium states. They are extended from a maximum-clique strategy and are global maximizers of the corresponding SPproblem. However, they are not strict maximizers, so they are only weakly evolutionarily stable. There are also non-clique strategies

that are not directly extended from a maximal clique strategy. For example, for the SNgame in Fig. 7.9, at equilibrium, the population is distributed on two disconnected cliques, with  $x_i^* = 1/8$  for all  $i = 1, \dots, 8$ . Note at this state, the only active constraint for the corresponding SPproblem is the equality constraint,  $\sum_i x_i = 1$ . A null space matrix  $Z$  for the Jacobian of this constraint can be defined as

$$Z = \begin{pmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then, it is easy to verify that

$$Z^T AZ = \begin{pmatrix} -2 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -2 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -2 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 0 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 0 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 0 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

Let  $u = (1, 1, 1, -1, -1, -1, -1)^T$ . Then,  $u^T Z^T AZ u = 24 > 0$ . Therefore,  $Z^T AZ$  cannot be negative semidefinite. By the second-order necessary condition for constrained optimization (107; 114),  $x^*$  cannot be a local maximizer of the SPproblem. It follows that  $x^*$  cannot even be weakly evolutionarily stable.

In general, a non-clique equilibrium state cannot be evolutionarily stable, and is at best a weakly evolutionarily stable state, as can be further justified in Theorem 7.4.12. On the other hand, equilibrium states on maximal cliques are evolutionarily stable except for some special circumstances, as we have shown in previous subsections. In this sense, the optimal clique strategies must have an evolutionary advantage over the non-clique optimal strategies.

**Theorem 7.4.12.** *An optimal strategy  $x^*$  on a non-clique subgraph  $H$  of  $G$  for the SNgame in (7.1) cannot be a strict local maximizer of the corresponding SPproblem in (7.4) and hence be evolutionarily stable.*

*Proof.* Let  $x^* \in S$  be an optimal strategy for the SNgame in (7.1),  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , where  $H$  is a non-clique subgraph of  $G$ . Assume that  $x^*$  is a local maximizer of the corresponding SPproblem in (7.4), i.e., there is a small neighborhood  $U$  of  $x^*$  such that  $x^{*T} Ax^* \geq y^T Ay$  for all  $y \in U \cap S$ . We show that  $x^*$  is not a strict local maximizer, i.e., there is no neighborhood  $U$  of  $x^*$  such that  $x^{*T} Ax^* > y^T Ay$  for all  $y \in U \cap S$ .

Since  $H$  is a non-clique subgraph, there must be  $i, j \in V_H$ ,  $(i, j) \notin E_H$ . Define  $p \in R^n$ , with  $p_i = \epsilon$ ,  $p_j = -\epsilon$ , and  $p_r = 0$  for all  $r \neq i, j$ . Then, by choosing  $\epsilon > 0$  sufficiently small in between  $-x_i^*$  and  $x_j^*$ , we can have  $y = x^* + p \in U \cap S$  for any neighborhood  $U$  of  $x^*$ . Note that

$$y^T Ay = (x^* + p)^T A(x^* + p) = x^{*T} Ax^* + 2x^{*T} Ap + p^T Ap.$$

Since  $x^*$  is a local maximizer of the SPproblem in (7.4), it must satisfy the KKT conditions of the SPproblem in (7.4). In particular, there must be a number  $\lambda^*$  such that  $(Ax^*)_r = \lambda^*$  for all  $r \in V_H$ . It follows that  $2x^{*T} Ap = 2\lambda^*(\epsilon - \epsilon) = 0$ . Since  $(i, j) \notin E_H$ ,  $A_{i,j} = A_{j,i} = 0$ , and  $p^T Ap = 0$ . It follows that  $y^T Ay = x^{*T} Ax^*$ . This simply implies that we cannot find any small neighborhood  $U$  of  $x^*$  such that  $x^{*T} Ax^* > y^T Ay$  for all  $y \in U \cap S$ . Therefore,  $x^*$  cannot be a strict local maximizer of the SPproblem and hence be evolutionarily stable.  $\square$

## 7.5 Conclusion

The optimal strategies or in other words, the equilibrium states of a social network game tend to be developed around the network cliques or their extensions. Not all cliques are optimal choices for a game, however. Only maximal cliques or some of their extensions may be selected at equilibrium. Yet, these optimal strategies may or may not be evolutionarily stable. There are non-clique equilibrium states as well, but they are certainly evolutionarily unstable. In this paper, we have conducted a detailed analysis on the clique and non-clique strategies of the social network games. We have in particular found the optimality and stability conditions on these strategies and provided rigorous mathematical justifications.

A social network game is a symmetric evolutionary game and therefore, corresponds to a generalized knapsack problem. An optimal strategy for the social network game is a KKT point for the corresponding generalized knapsack problem, which could be a local minimizer, saddle point, local

maximizer, or strict local maximizer. The strategy is evolutionarily unstable if it is a local minimizer or saddle point of the generalized knapsack problem, and is evolutionarily stable if it is a strict local maximizer. It is weakly evolutionarily stable if it is a non-strict local maximizer. These properties form the theoretical basis for analyzing the optimal strategies of the social network games. In non-mathematical terms, the results from our analysis can be summarized as follows.

- First, we have shown that the strategy on an arbitrary clique will not be optimal if the clique is contained in a larger clique. In general, the population on a social network would extend from a smaller clique to a larger one whenever possible, for achieving a higher social payoff (Fig. 7.3).
- The strategy on a maximal clique is always optimal, when the population is distributed evenly over the nodes of the clique. A maximal clique is not necessarily a maximum clique, but still, the network can reach equilibrium on such a clique (Fig. 7.4).
- Not every optimal strategy is on a maximal clique. A non-clique optimal strategy can be obtained when a maximal clique can be expanded to a certain larger subgraph (Fig. 7.5).
- As an optimal strategy, a maximal clique strategy is only a KKT point for the generalized knapsack problem. It may be a local minimizer, saddle point, local maximizer, or strict local maximizer, each with a different stability condition.
- When a larger clique is attached to  $k - 1$  nodes of a maximal clique of size  $k$ , the strategy on this maximal clique cannot even be a local maximizer for the corresponding generalized knapsack problem. It is therefore evolutionarily unstable. In this situation, the population, although at equilibrium, will still be able to extend from this maximal clique to its neighboring larger clique and then achieve a higher social payoff (Fig. 7.6).
- If there is no larger attached clique, the maximal clique strategy corresponds to a local maximizer of the generalized knapsack problem, and it is at least weakly evolutionarily stable.
- If there is still a clique of the same size attached to  $k - 1$  nodes of the maximal clique, the strategy on this maximal clique is not a strict local maximizer. Therefore, it can only be weakly evolutionarily stable. In this situation, the population, although at equilibrium, will still be able to

extend from this maximal clique to its neighboring clique without losing any social payoff (Fig. 7.7).

- If there is no attached clique of the same size, the maximal clique strategy corresponds to a strict local maximizer of the generalized knapsack problem, and is evolutionarily stable. This is a stable state of the network, when the population remains on the clique, and any change would reduce the social payoff.
- The strategy on a maximum clique is certainly an optimal strategy and also has the maximum payoff. However, it is a global maximizer of the generalized knapsack problem, which is computationally hard to obtain. Therefore, a maximum clique strategy may not be found easily in nature as well.
- The strategy on a maximum clique is at least weakly evolutionarily stable, for it always corresponds to a global maximizer of the generalized knapsack problem. It is evolutionarily stable when it is a strict global maximizer.
- When it is not a strict global maximizer, there is a procedure to change from this clique strategy to a larger non-clique one, and vice versa. In this situation, the population can spread its distribution from the maximum clique to a larger subgraph that contains the clique or vice versa (Fig. 7.8).
- A non-clique optimal strategy is evolutionarily unstable in general, and is at best weakly evolutionarily stable (Fig. 7.9). The optimal cliques strategies are evolutionarily stable except for special circumstances, and therefore, should have an evolutionary advantage over the non-clique ones.

The study on social cliques has a long history in network science and especially in social network analysis (109; 112; 117; 119). The maximum clique problem has been a classical subject in graph theory and combinatorial optimization (104), where the work was focused on the computation of the maximum clique of a given graph. The problem has been formulated and solved as a generalized knapsack problem in (113; 115). The relationship between the symmetric evolutionary game (SEgame) and the generalized knapsack problem (GKproblem) has been recognized in evolutionary game theory

in early 1980s, and the equivalence of the evolutionary stable state of the SEgame and the strict local maximizer of the GKproblem was then established (110; 116; 120). Later, the clique problem has been studied as a symmetric evolutionary game or a generalized knapsack problem in (106; 118), where the number of stable equilibria for the corresponding SEgame and the number of local maxima of the corresponding GKproblem were analyzed. Along this line, we have in this paper considered the clique problem as a social network game, and performed an in-depth analysis on the equilibrium states of the game. We have analyzed several different types of equilibrium states and proved a set of conditions for their stabilities. We have shown in particular that the equilibrium states on cliques are evolutionarily stable except for special circumstances, while non-clique equilibrium states are unstable in general. Therefore, the equilibrium states on cliques should have an evolutionary advantage over the non-clique ones. The results from our analysis are believed to be interesting for the understanding of the behaviors of social networks of either biological, or cultural, or economic types.

### **Acknowledgments**

This work is partially supported by the NIH/NIGMS grants R01GM-072014 and R01GM081680 and by the NSF/DMS grant DMS0914354. The authors would like to thank Professor Yaxiang Yuan, Professor Yuhong Dai, and Professor Xin Liu of Chinese Academy of Sciences for many insightful discussions in the early stage of this work. The authors would also like to thank Professor Stephen Wright of University of Wisconsin at Madison for taking time to read the early draft of the paper and providing valuable suggestions.

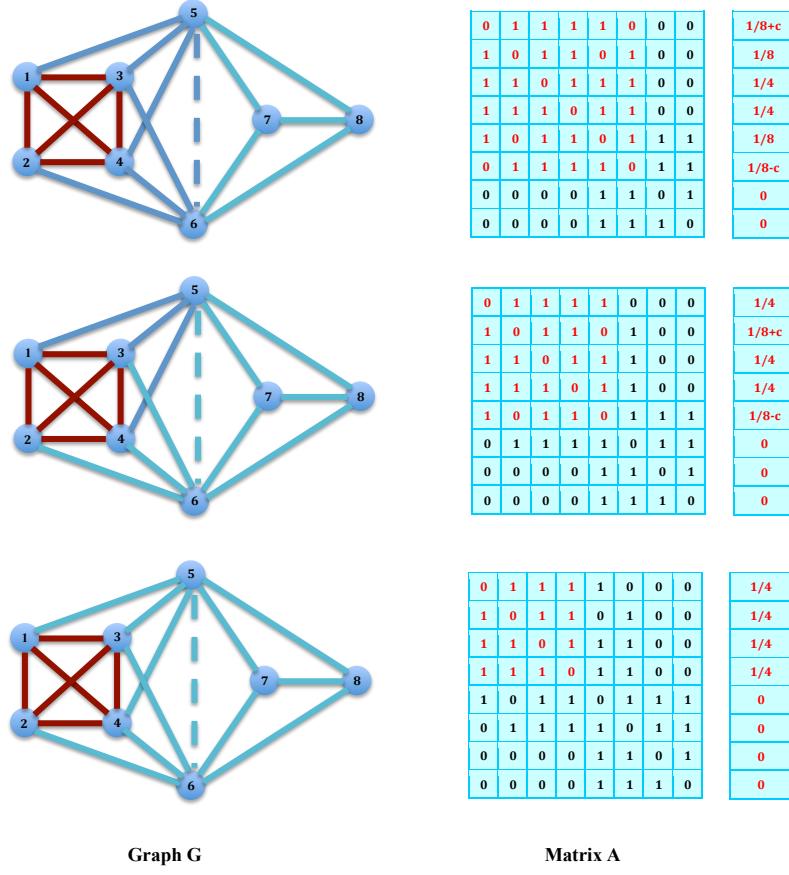


Figure 7.8 Recovering Maximum Cliques: In the graph on top, the nodes  $\{1, 2, 3, 4, 5, 6\}$  form a subgraph  $H$ , and  $x^* \in S$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , is a global maximizer for the SPproblem. However,  $H$  is not a maximum clique. Since node 2 and 6 are not connected, we therefore add  $1/8$  to  $x_2$  but subtract it from  $x_6$ . We then obtain a reduced subgraph  $H$ , with  $V_H = \{1, 2, 3, 4, 5\}$ ,  $E_H = \{(i, j) \in E : i, j \in V_H\}$ , as shown in the second graph to the top. The solution  $x^* \in S$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , remains to be a global maximizer for the SPproblem. Next, since node 1 and 5 are still not connected, we then add  $1/8$  to  $x_1$  but subtract it from  $x_5$  to obtain a further reduced subgraph  $H = (V_H, E_H)$ ,  $V_H = \{1, 2, 3, 4\}$ ,  $E_H = \{(i, j) : i, j \in V_H\}$ , as shown in the graph in the bottom. The solution  $x^* \in S$ ,  $x_i^* > 0$  for all  $i \in V_H$  and  $x_i^* = 0$  for all  $i \in V \setminus V_H$ , again remains to be a global maximizer for the SPproblem, but this time,  $H$  is a maximum clique of  $G$ .

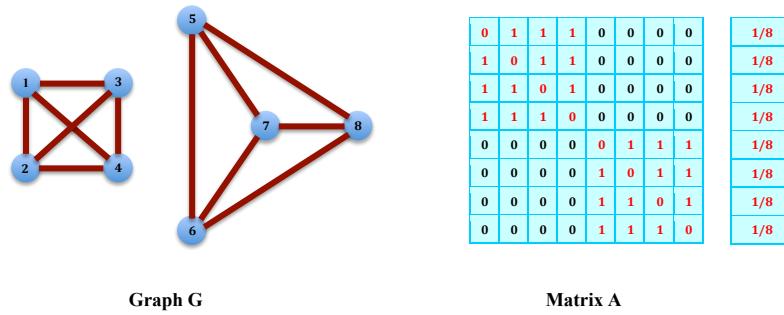


Figure 7.9 Non-Clique Equilibrium States: One of the equilibrium states of the SNgame on graph  $G$  in the figure is when the population is distributed on the two disconnected cliques, with  $x_i^* = 1/8$  for all  $i = 1, \dots, 8$ . It is easy to verify that this strategy is not a local maximizer of the corresponding SPproblem and thereby, is not even weakly evolutionarily stable.

## CHAPTER 8. SUMMARY AND CONCLUSION

We study the spatial effects on evolution of models, optimality and stability of symmetric game and its applications, and parameter estimation in game theoretic modeling of the biological and social systems. In summary, we categorize the types of nonlinear games and investigate the simulation study of spatial effects on the evolution of cooperation in a nonlinear yeast game. We also incorporate the spatial diffusion effects in the replicator dynamics of nonlinear game models, and prove the asymptotic behavior of the solution to the corresponding adapted replicator diffusion equation. Furthermore, we apply the statistical techniques and methodologies to solve the inverse problem in evolutionary game dynamics, that is, the NB mixture model and Markov random field model incorporating replicator equations are built where penalized approximated negative log-likelihood method with generalized smoothing approach and Besag pseudo-likelihood method are implemented to facilitate the estimation and inference of model parameters. Finally, the theory for obtaining optimal and stable strategies for symmetric evolutionary games is explored, and new proofs and computational methods are provided. Also the symmetric evolutionary game is applied to model the evolution of a population over a social network, and several different types of equilibrium states corresponding to social cliques are analyzed with a set of conditions for their stabilities proved.

### 8.1 2D Yeast Cooperation Game

Evolutionary game theory has achieved tremendous success in modeling biological systems. However almost all the successes have been focus on the linear games. The first applicable nonlinear game model is proposed by Gore et al. in 2009 and later the mathematical theory, especially the theoretical properties of this model is developed by us. We now make a further investigation of this nonlinear game from the perspective of 2D simulation and the contributions of this project can be listed as follows.

First of all, we have reviewed and generalized three types of linear games whose definitions are based on the comparative relationship of the elements in the so-called payoff matrix. The generalized nonlinear game are defined with nonlinear payoff functions instead of matrix-based linear functions. The Prisoner's Dilemma and Snowdrift game for nonlinear games are conceptualized from the perspective of payoff functions' curves and intercepts of payoff functions which are comparably same to those in matrix games, except the skewness and concavity of the curves. Parameters in payoff matrix as now replaced by environmental or internal parameters in nonlinear payoff functions. We also study the replicator dynamics of nonlinear games and would expect the nonlinear games can be widely use for a large group of systems.

Secondly, the combination of numerical simulation and evolutionary game theory has been proved to be useful to study the linear game models and how the spatial effects impact the evolution of co-operation. We extend the combined model to the field of nonlinear game models where rare scientist has considered before. We find that the spatial structure on a  $2D$  lattice tends to promote the evolution of cooperation in both Prisoner's Dilemma and Snowdrift Game, which is a little different from those found in matrix games since Christoph shows that spatial effects increase the proportion of cooperation in Prisoner's Dilemma while this doesn't hold for Snowdrift Game in matrix games. We also compare the deterministic and stochastic update rules and conclude that their results are different from each other but tend to coincide as neighborhood size increases.

Thirdly, we propose the difference equations for the  $2D$  simulation that are implemented in Section 2.5. Then we analyze some basic properties of the difference equations for the systems of players. The invasion conditions are provided given four structures for two different situations with the neighborhood size  $N = 4$ . We find some interesting patterns for the inequalities of the invasion conditions under different combinations of parameters from the experimental data Gore et.al collected and those patterns convince us the promotion of cooperation in the numerical experiments. We also make a generalization based on neighborhood sizes and update rules and concluding results are discussed.

Finally, this project invites more interest from both mathematical and biological communities in evolutionary game theory and its applications. It's an extension of the work about the game dynamic model for yeast system proposed by Huang and Wu (2012). Further development of our work can be fascinating, such as the rigorous theoretical properties for the  $2D$  model on the lattice and so on.

## 8.2 ESS in Nonlinear Game and the Asymptotic Behavior of Solution to Adapted Replicator Diffusion Equation

In this project, we have discussed the equivalent conditions of the evolutionarily stable strategy for a general class of replicator dynamics. We also establish the theory for the long time behavior of the solution in the adapted replicator diffusion equations which incorporates the spatial effects. Finally the construct artificial system and the yeast cooperation system in Gore et al.(2009) are tested out to verify the long time behavior of the adapted replicator diffusion equations generated from the game dynamics.

Although it's difficult to find the necessary and sufficient conditions of the existence of evolutionarily stable strategies for the direct form of the payoff functions of all species in a population game based on the replicator dynamics, we derive the equivalent conditions of ESS from the differential behavior of the payoff functions at the stationary state indirectly using the Taylor expansion idea. The coefficients of two Taylor series are compared based on the definition of the ESS and this yields to infinite series of inequalities involving the infinite differential properties of the payoff functions at the stationary state. We find that the stronger version of first two inequalities are necessary to the existence of a regular ESS which is an refinement of ESS. Also these two inequalities are necessary and sufficient for an ESS when the game is a linear game.

As for the general replicator dynamics of multiple species in a population, the spatial effects determined by the environment are introduced and the adapted replicator diffusion equation are then generated from the general replicator equations by adding the same diffusion term for all species. Based on the lemmas about the convergence of simple heat equation and the existence and uniqueness of a general class of parabolic equations, we prove the uniform convergence of the solution of the adapted replicator diffusion equation to the exponentially stable stationary state. This suggests that the cooperation and competition of several species in a spatially distributed population finally reaches an equilibrium which is the same as those well mixed population in the same payoff functions situation from the biological point of view.

The constructed game and the yeast game proposed by Gore et al. are classical population games within two populations. In the yeast game, Gore et al. examined the wild type strain (cooperator) and the mutant strain (defector). The authors derived the payoff functions for this yeast system which was

a nonlinear game. The dynamics and equilibrium stability of this game were investigated by Huang and Wu (2012) and the results fitted the experimental ones very well. The replicator equations for the artificial game, and especially for the yeast game, are adapted by adding diffusion terms because of the nature of strains digesting the dispersal glucose in the environment. And we use the semi-implicit algorithm to generate the numerical solutions of the two adapted equations. The contour plots of the frequencies of the strategy  $C$  are plotted, and they indicates that it is the diffusion effect that drives the behaviors of the solutions initially. The latter process are then governed by the replicator source term and this is justified by the contour plots of cooperator frequencies and the comparison plots of the frequencies under the well-mixed replicator dynamics and the average frequencies under the adapted replicator diffusion equations for different types of nonlinear games. The payoff functions in yeast game don't satisfy both conditions in theorem 3.3.5, but the numerical experiments still indicates the long time behavior of the solution, which suggests that the conditions in theorem 3.3.5 may be stronger for the asymptotic convergence of the solution.

The present modeling study shows the asymptotical behavior of the adapted replicator diffusion equations under the same diffusion rate situation. It also filling one with the interest in evolutionary game theory and its applicators from both mathematical and biological perspective. Our future work will consider the behavior of the solution to the parabolic systems with weaker conditions than those in theorem 3.3.5 and more general diffusion effects which means the dispersal rate of the species in a population vary or be more general operator rather than Laplacian, and the methodology of parameter estimation in such replicator diffusion equations.

### 8.3 Parameter Estimation in Dynamical Metagenomic Model

In this project, we explore the use of the replicator equations for studying the interaction of species in a microbial community sample over time using next generation sequencing applied to metagenomics. Specifically, we propose a parsimonious mixture model to reduce the number of free parameters. We combine information across multiple biological replicates, using the Negative Binomial to account for excess biological variation. We use smoothing spline approximation to the replicator equation solu-

tions to avoid repeated integration. Finally, we demonstrate the feasibility of our model by testing it on simulation data. We end by discussing future directions that will further improve our estimation procedure.

The estimation procedure is so far quite slow, so we were only able to perform one complete EM run. Thus it was impossible to assess what might have failed or what could be improved in the method. For the future, it will be helpful to rerun the estimation while holding some subsets of the parameters at their true values. In this way, we can assess whether the observed misestimations are significantly impacting the estimation of  $\mathbf{b}$ , our main parameters of interest. In addition, we chose to use a smoothing spline approximation with 164 parameters. We should rerun the analysis with more or fewer smoothing spline parameters.

We used simulated data to assess our method for estimating interaction effects between species in mixed populations. We assumed the species interacted through replicator dynamics that hypothesize a specific payoff for interactions between pairs of species. Cooperating species  $i$  and  $j$  will receive positive payoffs,  $a_{ij}, a_{ji} > 0$ , from their interaction, while a parasitic species  $i$  will negatively impact its victim  $j$ ,  $a_{ji} < 0 < a_{ij}$ . Other interactions are also possible. In modern metagenomics datasets, there are many potentially interacting species, which leads to an explosion in the number of parameters. However, it is often true that many distinguishable species will interact similarly, so it is plausible to cluster species in homogeneously reacting groups. Our proposed model allows such grouping. Note, our grouping is partly genetic (reads are clustered into OTUs) and partly phenotypic (based on temporal dynamics). We feel this approach is better than pure genetic grouping, which treats species as less behaviorally plastic than they are. Future extensions could cluster genomes based on both criteria.

It is difficult to assess the performance of our method as the dataset size increases. Unfortunately, adding new data in any direction, more OTUs, more time points, or more mice, leads to concomitant increases in the number of parameters. The only way to increase the amount of data without adding more parameters is to increase  $s_{jm}$ . Of course, this also increases the variance. In real data applications, it may be possible to regularize  $s_{jm}$  across times or mice to control the number of parameters. Multiplex sequences may also help, and it may also work reasonably well to set  $s_{jm} = S_{jm}$ .

Another completely different approach is to dispense with the smoothing spline approximation. Instead, we can utilize a Newton-Raphson optimization, which should converge in fewer iterations,

thus requiring relatively few numerical integrations of the replication equations and similar equations for the Hessians. Fortunately, for the cluster model, the payoff matrix is relatively small.

#### **8.4 Markov Random Field Model for Analysis of Game Dynamics**

Inverse problem is a well-studied mathematical problem and arise in many branches of science and mathematics. The conceptual understanding is that the transformation from data to model parameters is a result of the interaction of a physical system with the object that we wish to infer about. To estimate the parameters of the payoff matrix for linear game model, we introduce the main idea on development of Markov Random Field Model, which incorporates replicator dynamics and mainly deals with observations on a lattice using appropriate random variable distribution, to estimate the parameters of the payoff matrix for linear game model. In the near future, a simulation study will be explored and study of different games with respect to planes consist of paired parameters will be implemented. And the parameter estimation procedure will be a process of maximization of the Besag likelihood function over the parameter space for the inverse spatial game, and then the inference of the parameters will be based a MCMC method. Furthermore, model assessment approaches will be followed to justify the performance of the proposed Markov Random Field model. Finally we will investigate the spatial effects of interaction region or neighborhood size on the estimated parameters for given data and analyze the nature of the game that we are interested in. It is expected that the Markov Random Field model will be useful in applying to a lot of spatial-temporal studies consist of two species or subpopulations in biological or social systems.

#### **8.5 Optimality and Stability of Symmetric Evolutionary Games with Applications in Genetic Selection**

In this project, we have reviewed the theory for obtaining optimal and stable strategies for SEgames, and provided some new proofs and computational methods. In particular, we have reviewed the relationship between the SEgame and the GKproblem, and discussed the first and second order necessary and sufficient conditions that can be derived from this relationship for testing the optimality and stability of the strategies. Some of the conditions are given in different forms from those in previous work

and can be verified more efficiently. We have also derived more efficient computational methods for the evaluation of the conditions than conventional approaches. We have demonstrated how these conditions can be applied to justifying the strategies and their stabilities for a special class of genetic selection games including some in the study of genetic disorders. Further studies can be pursued in the following possible directions though:

First, novel methods can be developed for solving special types of SEgames and especially for obtaining the evolutionarily stable strategies for the games by solving some special classes of GKproblems. For example, if the fitness matrix for a SEgame is negative definite, then the corresponding GKproblem is a strictly convex quadratic program and can be solved efficiently using some special algorithms. Further, the solution is guaranteed to be a strict local maximizer for the GKproblem and hence an evolutionarily stable strategy for the SEgame. A more complicated case is when the fitness matrix is positive definite. Then, only pure strategies may be evolutionarily stable. A special algorithm can then be developed to only find the solutions for the GKproblem that correspond to the pure strategies of the SEgame.

Second, in Theorem 6.3.5 and 6.5.3, we have stated two optimality and stability conditions. They are necessary and sufficient, but require all active constraints to be strongly active at  $x^*$ , when  $C_0(x^*) = C^0(x^*)$ ,  $T_0(x^*) = T^0(x^*)$ , and  $Z_0 = Z^0$ . However, in practice, this assumption may not hold. A more general necessary and sufficient condition, without the above assumption, is to require  $d^T Ad < 0$  for all  $d \in T(x^*)$ ,  $d \neq 0$ , where  $T(x^*)$  is the reduced tangent cone at  $x^*$ , as given in Bomze's paper(2002). As we have mentioned in previous sections, this condition is not easy to test. It is equivalent to testing the copositivity of a matrix, which is difficult in general. But still, an efficient algorithm may be developed for SEgames and GKproblems for small sizes of problems or problems with special structures.

Third, it is not so hard to verify that the GKproblem is NP-hard in general, because the maximum clique problem can be formulated as a GKproblem. However, how to extend this result to the SEgame is not so clear, because the SEgame is not exactly equivalent to the GKproblem. Several related questions are asked: is any maximal clique a local maximizer of the GKproblem for the maximum clique problem? If not, what condition is needed? If yes, is it a strict local maximizer? Is the maximum clique a global maximizer? Is it an evolutionarily stable strategy for the corresponding SEgame? We are interested in all these questions and are trying to find their answers.

Fourth, though not equivalent, the correspondence between the SEgame and GKproblem is interesting. A similar relationship may be found between a class of nonlinear games and nonlinear optimization problems. Indeed, we can define an n-strategy two-player game by a fitness function  $x^T \pi(y)$  with  $\pi(y)$  being a nonlinear function. The game then becomes a nonlinear game. If  $\pi(y)$  is a gradient field, i.e., there is a function  $f(y)$  such that  $\nabla f(y) = \pi(y)$ , then, an optimal strategy  $x^* \in S$  such that  $x^{*T} \pi(x^*) \geq x^T \pi(x^*)$  for all  $x \in S$  corresponds to an optimal solution  $x^* \in S$  such that  $f(x^*) \geq f(x)$  for all  $x$  in a small neighborhood of  $x^*$ ,  $x \in S$ . Then, it would be interesting to see what additional relationships between the SEgame and GKproblem can be extended to their nonlinear cases.

Finally, we have demonstrated the applications of SEgames to allele selection at single genetic loci. They can be extended to alleles at multiple genetic loci, if there is no mutation or recombination. In this case, an individual can be identified by a sequence of alleles at the multiple loci. In other words, a selection strategy will be a choice of a specific sequence of alleles. This would certainly increase the strategy space substantially. For example, if there are two loci  $G_1$  and  $G_2$ , with two possible alleles  $A$  and  $a$  for  $G_1$  and two other possible ones  $B$  and  $b$  for  $G_2$ , then there will be four possible sequences of alleles for the two loci:  $AB$ ,  $Ab$ ,  $aB$ ,  $ab$ , each corresponding to one pure strategy. In general, if there are  $m$  loci  $G_i$ ,  $i = 1, \dots, m$ , with  $m_i$  possible alleles for  $G_i$ , then there will be  $n = \prod_{i=1:m} m_i$  possible sequences of alleles. The number of pure strategies and hence the dimension of the game will be  $n$ , which can be a large number. In any case, in practice, mutation and recombination often are not negligible, and therefore, our model must incorporate such effects. The topics could include other so-called linkage disequilibrium factors, but they are all beyond the scope of this project. We will pursue these issues in our future efforts.

## 8.6 Optimal Strategies for Social Network Games

The optimal strategies or in other words, the equilibrium states of a social network game tend to be developed around the network cliques or their extensions. Not all cliques are optimal choices for a game, however. Only maximal cliques or some of their extensions may be selected at equilibrium. Yet, these optimal strategies may or may not be evolutionarily stable. There are non-clique equilibrium states as well, but they are certainly evolutionarily unstable. In this project, we have conducted a detailed

analysis on the clique and non-clique strategies of the social network games. We have in particular found the optimality and stability conditions on these strategies and provided rigorous mathematical justifications.

A social network game is a symmetric evolutionary game and therefore, corresponds to a generalized knapsack problem. An optimal strategy for the social network game is a KKT point for the corresponding generalized knapsack problem, which could be a local minimizer, saddle point, local maximizer, or strict local maximizer. The strategy is evolutionarily unstable if it is a local minimizer or saddle point of the generalized knapsack problem, and is evolutionarily stable if it is a strict local maximizer. It is weakly evolutionarily stable if it is a non-strict local maximizer. These properties form the theoretical basis for analyzing the optimal strategies of the social network games. In non-mathematical terms, the results from our analysis can be summarized as follows.

- First, we have shown that the strategy on an arbitrary clique will not be optimal if the clique is contained in a larger clique. In general, the population on a social network would extend from a smaller clique to a larger one whenever possible, for achieving a higher social payoff (Fig. 7.3).
- The strategy on a maximal clique is always optimal, when the population is distributed evenly over the nodes of the clique. A maximal clique is not necessarily a maximum clique, but still, the network can reach equilibrium on such a clique (Fig. 7.4).
- Not every optimal strategy is on a maximal clique. A non-clique optimal strategy can be obtained when a maximal clique can be expanded to a certain larger subgraph (Fig. 7.5).
- As an optimal strategy, a maximal clique strategy is only a KKT point for the generalized knapsack problem. It may be a local minimizer, saddle point, local maximizer, or strict local maximizer, each with a different stability condition.
- When a larger clique is attached to  $k - 1$  nodes of a maximal clique of size  $k$ , the strategy on this maximal clique cannot even be a local maximizer for the corresponding generalized knapsack problem. It is therefore evolutionarily unstable. In this situation, the population, although at equilibrium, will still be able to extend from this maximal clique to its neighboring larger clique and then achieve a higher social payoff (Fig. 7.6).

- If there is no larger attached clique, the maximal clique strategy corresponds to a local maximizer of the generalized knapsack problem, and it is at least weakly evolutionarily stable.
- If there is still a clique of the same size attached to  $k - 1$  nodes of the maximal clique, the strategy on this maximal clique is not a strict local maximizer. Therefore, it can only be weakly evolutionarily stable. In this situation, the population, although at equilibrium, will still be able to extend from this maximal clique to its neighboring clique without losing any social payoff (Fig. 7.7).
- If there is no attached clique of the same size, the maximal clique strategy corresponds to a strict local maximizer of the generalized knapsack problem, and is evolutionarily stable. This is a stable state of the network, when the population remains on the clique, and any change would reduce the social payoff.
- The strategy on a maximum clique is certainly an optimal strategy and also has the maximum payoff. However, it is a global maximizer of the generalized knapsack problem, which is computationally hard to obtain. Therefore, a maximum clique strategy may not be found easily in nature as well.
- The strategy on a maximum clique is at least weakly evolutionarily stable, for it always corresponds to a global maximizer of the generalized knapsack problem. It is evolutionarily stable when it is a strict global maximizer.
- When it is not a strict global maximizer, there is a procedure to change from this clique strategy to a larger non-clique one, and vice versa. In this situation, the population can spread its distribution from the maximum clique to a larger subgraph that contains the clique or vice versa (Fig. 7.8).
- A non-clique optimal strategy is evolutionarily unstable in general, and is at best weakly evolutionarily stable (Fig. 7.9). The optimal cliques strategies are evolutionarily stable except for special circumstances, and therefore, should have an evolutionary advantage over the non-clique ones.

The study on social cliques has a long history in network science and especially in social network analysis. The maximum clique problem has been a classical subject in graph theory and combinatorial optimization, where the work was focused on the computation of the maximum clique of a given graph. The problem has been formulated and solved as a generalized knapsack problem. The relationship between the symmetric evolutionary game (SEgame) and the generalized knapsack problem (GKproblem) has been recognized in evolutionary game theory in early 1980s, and the equivalence of the evolutionary stable state of the SEgame and the strict local maximizer of the GKproblem was then established. Later, the clique problem has been studied as a symmetric evolutionary game or a generalized knapsack problem, where the number of stable equilibria for the corresponding SEgame and the number of local maxima of the corresponding GKproblem were analyzed. Along this line, we have in this project considered the clique problem as a social network game, and performed an in-depth analysis on the equilibrium states of the game. We have analyzed several different types of equilibrium states and proved a set of conditions for their stabilities. We have shown in particular that the equilibrium states on cliques are evolutionarily stable except for special circumstances, while non-clique equilibrium states are unstable in general. Therefore, the equilibrium states on cliques should have an evolutionary advantage over the non-clique ones. The results from our analysis are believed to be interesting for the understanding of the behaviors of social networks of either biological, or cultural, or economic types.

## APPENDIX A. PROOF OF THEOREMS

Here the proof of the theorem in Section 4.3.1 is provided.

PROOF OF THEOREM 1.1 Let  $y = \sum_{i=1}^{n-1} y_i$ . From the transformation  $x_i = y_i/(1+y)$  for  $i = 1, \dots, n-1$  and  $x_n = 1/(1+y)$ , we have

$$\begin{aligned}
\dot{y} &= \sum_{i=1}^{n-1} \dot{y}_i \\
&= \sum_{i=1}^{n-1} y_i (r_i + \sum_{j=1}^{n-1} b_{ij} y_j) \\
&= \sum_{i=1}^{n-1} y_i (a_{in} - a_{nn}) + \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} (a_{ij} - a_{nj}) y_i y_j \\
&= \sum_{i=1}^{n-1} a_{in} x_i (1+y) - a_{nn} y + \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} (a_{ij} - a_{nj}) x_i x_j (1+y)^2 \\
&= (1+y) \sum_{i=1}^{n-1} a_{in} x_i - a_{nn} y + (1+y)^2 \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j - (1+y)^2 \sum_{i=1}^{n-1} x_i \sum_{j=1}^{n-1} a_{nj} x_j \\
&= (1+y) \sum_{i=1}^{n-1} a_{in} x_i - a_{nn} y + (1+y)^2 \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j - y(1+y) \sum_{j=1}^{n-1} a_{nj} x_j.
\end{aligned}$$

Thus

$$\begin{aligned}
\left( \frac{1}{1+y} \right)' &= -\frac{\dot{y}}{(1+y)^2} \\
&= -\frac{1}{1+y} \sum_{i=1}^{n-1} a_{in} x_i + a_{nn} \frac{y}{(1+y)^2} - \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j + \frac{y}{1+y} \sum_{j=1}^{n-1} a_{nj} x_j.
\end{aligned}$$

For  $i = 1, \dots, n-1$ , we have

$$\begin{aligned}
\dot{x}_i &= \left( \frac{y_i}{1+y} \right)' \\
&= y_i \left( \frac{1}{1+y} \right)' + \frac{\dot{y}_i}{1+y} \\
&= y_i \left[ -\frac{1}{1+y} \sum_{i=1}^{n-1} a_{in} x_i + a_{nn} \frac{y}{(1+y)^2} - \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j + \frac{y}{1+y} \sum_{j=1}^{n-1} a_{nj} x_j \right] + \frac{1}{1+y} y_i (r_i + \sum_{j=1}^{n-1} b_{ij} y_j) \\
&= -x_i \sum_{i=1}^{n-1} a_{in} x_i + a_{nn} x_i \frac{y}{1+y} - (1+y) x_i \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j + y x_i \sum_{j=1}^{n-1} a_{nj} x_j + x_i (a_{in} - a_{nn}) + (1+y) \cdot \\
&\quad x_i \sum_{j=1}^{n-1} (a_{ij} - a_{nj}) x_j \\
&= x_i \left[ -\sum_{i=1}^{n-1} a_{in} x_i + \frac{a_{nn} y}{1+y} - (1+y) \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j + y \sum_{j=1}^{n-1} a_{nj} x_j + a_{in} - a_{nn} + (1+y) \sum_{j=1}^{n-1} (a_{ij} - \right. \\
&\quad \left. a_{nj}) x_j \right] \\
&= x_i \left[ -\sum_{i=1}^{n-1} a_{in} x_i + \frac{a_{nn} y}{1+y} - (1+y) \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j + y \sum_{j=1}^{n-1} a_{nj} x_j + a_{in} - a_{nn} + (1+y) \sum_{j=1}^{n-1} a_{ij} x_j - \right. \\
&\quad \left. (1+y) \sum_{j=1}^{n-1} a_{nj} x_j \right] \\
&= x_i \left[ (1+y) \sum_{j=1}^{n-1} a_{ij} x_j + a_{in} - (1+y) \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j - (1+y) \sum_{i=1}^{n-1} a_{in} x_i x_n - (1+y) a_{nn} x_n^2 - (1+y) \cdot \right. \\
&\quad \left. \sum_{j=1}^{n-1} a_{nj} x_j x_n \right] \\
&= x_i \left[ (1+y) \sum_{j=1}^n a_{ij} x_j - (1+y) \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \right] \\
&= (1+y) \left\{ x_i \left[ \sum_{j=1}^n a_{ij} x_j - \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \right] \right\}.
\end{aligned}$$

For  $i = n$ , we have

$$\begin{aligned}
\dot{x}_n &= \left( \frac{1}{1+y} \right)' \\
&= -\frac{1}{1+y} \sum_{i=1}^{n-1} a_{in} x_i + a_{nn} \frac{y}{(1+y)^2} - \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j + \frac{y}{1+y} \sum_{j=1}^{n-1} a_{nj} x_j \\
&= -x_n \sum_{i=1}^{n-1} a_{in} x_i + a_{nn} x_n \frac{y}{(1+y)} - x_n (1+y) \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j + y x_n \sum_{j=1}^{n-1} a_{nj} x_j \\
&= x_n \left[ -\sum_{i=1}^{n-1} a_{in} x_i + a_{nn} \frac{y}{(1+y)} - (1+y) \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j + y \sum_{j=1}^{n-1} a_{nj} x_j \right] \\
&= x_n \left[ -\sum_{i=1}^{n-1} a_{in} x_i + a_{nn} \frac{y}{(1+y)} - (1+y) \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j + y \sum_{j=1}^{n-1} a_{nj} x_j \right] \\
&= x_n \left[ (1+y) \sum_{j=1}^{n-1} a_{nj} x_j - \sum_{j=1}^{n-1} a_{nj} x_j + a_{nn} \frac{1}{1+y} [(1+y) - 1] - \sum_{i=1}^{n-1} a_{in} x_i - (1+y) \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j \right] \\
&= x_n \left[ (1+y) \sum_{j=1}^{n-1} a_{nj} x_j - (1+y) \sum_{j=1}^{n-1} a_{nj} x_j x_n + (1+y) a_{nn} x_n - (1+y) a_{nn} x_n^2 - (1+y) \sum_{i=1}^{n-1} a_{in} x_i x_n - \right. \\
&\quad \left. (1+y) \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j \right] \\
&= x_n \left[ (1+y) \sum_{j=1}^n a_{nj} x_j - (1+y) \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \right] \\
&= (1+y) \left\{ x_n \left[ \sum_{j=1}^n a_{nj} x_j - \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \right] \right\}.
\end{aligned}$$

After the transformation, we obtain

$$\dot{x}_i = (1+y) \left\{ x_i \left[ \sum_{j=1}^n a_{ij} x_j - \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \right] \right\} \quad i = 1, \dots, n, \quad (\text{A.1})$$

which are equivalent to the replicator equations.

## APPENDIX B. LOGIT TRANSFORMATION IN EM ALGORITHM

In EM algorithm of Chapter 4, the expression of  $\hat{\mathbf{p}}^{(t+1)}$  is formulated. However, there are no explicit expressions for the formulas of  $(\hat{\mathbf{s}}^{(t+1)}, \hat{\mathbf{d}}^{(t+1)}, \hat{\mathbf{b}}^{(t+1)}, \hat{\mathbf{c}}^{(t+1)})$ , and we can use numerical optimization algorithms to compute them. That is, the first step of the EM algorithm can be implemented by trust region reflective method within the constrained optimization toolbox, and the second step can be achieved by the generalized profiling estimation procedure—iterative inner and outer criterion optimization algorithm. In the second step of the EM algorithm, the inner criteria is

$$\begin{aligned} H(\hat{\mathbf{c}}(\hat{\mathbf{b}})) = & - \sum_{k_1=1}^K \cdots \sum_{k_n=1}^K p_{1,k_1}^{(t)} \cdots p_{n,k_n}^{(t)} \left\{ (\log p_{k_1}^{(t+1)} + \cdots + \log p_{k_n}^{(t+1)}) + \sum_{i=1}^n \sum_{j=1}^N \sum_{m=1}^M [\log \Gamma(y_{ijm} + \hat{\phi}_{ij}^{-1}) \right. \\ & - \log \Gamma(\hat{\phi}_{ij}^{-1}) - \Gamma(y_{ijm} + 1) - (\hat{\phi}_{ij}^{-1} + y_{ijm}) \log(1 + s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} w_{k_i,j}(\hat{\mathbf{c}}) \hat{\phi}_{ij}) + y_{ijm} \log(s_{jm}^{(t+1)} \cdot \\ & \left. d_{i,k_i}^{(t+1)} w_{k_i,j}(\hat{\mathbf{c}}) \hat{\phi}_{ij})] \right\}. \end{aligned}$$

And the outer criteria is

$$\begin{aligned} J(\hat{\mathbf{c}}, \hat{\mathbf{b}} | \lambda) = & - \sum_{k_1=1}^K \cdots \sum_{k_n=1}^K p_{1,k_1}^{(t)} \cdots p_{n,k_n}^{(t)} \left\{ (\log p_{k_1}^{(t+1)} + \cdots + \log p_{k_n}^{(t+1)}) + \sum_{i=1}^n \sum_{j=1}^N \sum_{m=1}^M [\log \Gamma(y_{ijm} + \hat{\phi}_{ij}^{-1}) \right. \\ & - \log \Gamma(\hat{\phi}_{ij}^{-1}) - \Gamma(y_{ijm} + 1) - (\hat{\phi}_{ij}^{-1} + y_{ijm}) \log(1 + s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} w_{k_i,j}(\hat{\mathbf{c}}) \hat{\phi}_{ij}) + y_{ijm} \log(s_{jm}^{(t+1)} \cdot \\ & \left. d_{i,k_i}^{(t+1)} w_{k_i,j}(\hat{\mathbf{c}}) \hat{\phi}_{ij})] \right\} + \lambda \sum_{m=1}^{K-1} \int_{[t_1, t_N]} \left[ \frac{dw_m(t|\mathbf{c})}{dt} - g_m(\mathbf{w}(t|\hat{\mathbf{c}}), \hat{\mathbf{b}}) \right]^2 dt, \end{aligned}$$

where  $\mathbf{w}(t)$  will be described in later sessions. The partial derivatives of  $J(\hat{\mathbf{c}}, \hat{\mathbf{b}} | \lambda)$  with respect to  $c_{kp}$ ,  $k = 1, \dots, K-1; p = 1, \dots, M$  are

$$\begin{aligned} \frac{\partial J}{\partial c_{kp}} = & - \sum_{k_1=1}^K \cdots \sum_{k_n=1}^K p_{1,k_1}^{(t)} \cdots p_{n,k_n}^{(t)} \left\{ \sum_{i=1}^n \sum_{j=1}^N \sum_{m=1}^M \left[ -(\hat{\phi}_{ij}^{-1} + y_{ijm}) \frac{s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} \hat{\phi}_{ij}}{1 + s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} w_{k_i,j}(\hat{\mathbf{c}}) \hat{\phi}_{ij}} \frac{\partial w_{k_i,j}(\hat{\mathbf{c}})}{\partial c_{kp}} + y_{ijm} \cdot \right. \right. \\ & \left. \left. \frac{1}{w_{k_i,j}(\hat{\mathbf{c}})} \frac{\partial w_{k_i,j}(\hat{\mathbf{c}})}{\partial c_{kp}} \right] \right\} + 2\lambda \sum_{m=1}^{K-1} \int_{[t_1, t_N]} \left[ \dot{w}_m(t|\hat{\mathbf{c}}) - g_m(\mathbf{w}(t|\hat{\mathbf{c}}), \hat{\mathbf{b}}) \right] \left[ \frac{\partial \dot{w}_m(t|\mathbf{c})}{\partial c_{kp}} - \sum_{s=1}^{K-1} \frac{\partial g_m}{\partial w_s} \frac{\partial w_s}{\partial c_{kp}} \right] dt. \end{aligned}$$

And the second partial derivatives of  $J(\underline{c}, \underline{b}|\lambda)$  with respect to  $\underline{c}$  are

$$\begin{aligned} \frac{\partial^2 J}{\partial c_{kp} \partial c_{lq}} = & - \sum_{k_1=1}^K \cdots \sum_{k_n=1}^K p_{1,k_1}^{(t)} \cdots p_{n,k_n}^{(t)} \left\{ \sum_{i=1}^n \sum_{j=1}^N \sum_{m=1}^M \left[ (\hat{\phi}_{ij}^{-1} + y_{ijm}) \frac{(s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} \hat{\phi}_{ij})^2}{(1+s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} w_{k_i,j}(\underline{c}) \hat{\phi}_{ij})^2} \frac{\partial w_{k_i,j}(\underline{c})}{\partial c_{kp}} \right. \right. \\ & \left. \frac{\partial w_{k_i,j}(\underline{c})}{\partial c_{lq}} - (\hat{\phi}_{ij}^{-1} + y_{ijm}) \frac{s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} \hat{\phi}_{ij}}{1+s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} w_{k_i,j}(\underline{c}) \hat{\phi}_{ij}} \frac{\partial^2 w_{k_i,j}(\underline{c})}{\partial c_{kp} \partial c_{lq}} - y_{ijm} \frac{1}{w_{k_i,j}(\underline{c})^2} \frac{\partial w_{k_i,j}(\underline{c})}{\partial c_{kp}} \frac{\partial w_{k_i,j}(\underline{c})}{\partial c_{lq}} + \right. \\ & \left. \left. y_{ijm} \frac{1}{w_{k_i,j}(\underline{c})} \frac{\partial^2 w_{k_i,j}(\underline{c})}{\partial c_{kp} \partial c_{lq}} \right] \right\} + 2\lambda \sum_{m=1}^{K-1} \int_{[t_1, t_N]} \left\{ \left[ \frac{\partial \dot{w}_m(t|\underline{c})}{\partial c_{lq}} - \sum_{s=1}^{K-1} \frac{\partial g_m}{\partial w_s} \frac{\partial w_s}{\partial c_{lq}} \right] \left[ \frac{\partial \dot{w}_m(t|\underline{c})}{\partial c_{kp}} - \sum_{s=1}^{K-1} \frac{\partial g_m}{\partial w_s} \right] \right. \\ & \left. \left. \frac{\partial w_s}{\partial c_{kp}} \right] + [\dot{w}_m(t|\underline{c}) - g_m(\underline{w}(t|\underline{c}, \underline{p}))] \left[ \frac{\partial^2 \dot{w}_m(t|\underline{c})}{\partial c_{kp} \partial c_{lq}} - \sum_{s=1}^{K-1} \left( \frac{\partial w_s}{\partial c_{kp}} \sum_{r=1}^{K-1} \frac{\partial^2 g_m}{\partial w_s \partial w_r} \frac{\partial w_r}{\partial c_{lq}} + \frac{\partial g_m}{\partial w_s} \frac{\partial^2 w_s}{\partial c_{kp} \partial c_{lq}} \right) \right] \right\} dt. \end{aligned}$$

The second partial derivatives of  $J(\underline{c}, \underline{b}|\lambda)$  with respect to  $\underline{c}$  and  $\underline{b}$  are

$$\frac{\partial^2 J}{\partial c_{kp} \partial b_z} = 2\lambda \sum_{m=1}^{K-1} \int_{[t_1, t_N]} \left[ - \frac{\partial g_m}{\partial b_z} \left( \frac{\partial \dot{w}_m(t|\underline{c})}{\partial c_{kp}} - \sum_{s=1}^{K-1} \frac{\partial g_m}{\partial w_s} \frac{\partial w_s}{\partial c_{kp}} \right) - \sum_{s=1}^{K-1} \frac{\partial^2 g_m}{\partial w_s \partial b_z} \frac{\partial w_s}{\partial c_{kp}} (\dot{w}_m - g_m(\underline{w})) \right] dt.$$

In the model description part, the frequencies can be approximated by smoothing splines (polynomials, trigonometric functions, etc.). However, the frequencies are always non-negative while the direct applications of those splines may yield to negative approximates. To tackle this technical problem, we introduce the logistic transformation (one minor issue is that the smoothing splines could be very close but not equal to 0 when the true frequencies are 0 at some time sub-interval):

$$\begin{aligned} \log \frac{w_1}{w_K} &= \sum_{j=1}^M c_{1j} \phi_{1j}(t) \\ &\vdots \\ \log \frac{w_{K-1}}{w_K} &= \sum_{j=1}^M c_{K-1,j} \phi_{K-1,j}(t) \\ \log \frac{w_K}{w_K} &= 0. \end{aligned}$$

That is,

$$\begin{aligned}
 w_1 &= \frac{\exp\left(\sum_{j=1}^M c_{1j} \phi_{1j}(t)\right)}{1 + \sum_{k=1}^{K-1} \exp\left(\sum_{j=1}^M c_{kj} \phi_{kj}(t)\right)} \\
 &\quad \vdots \\
 w_{K-1} &= \frac{\exp\left(\sum_{j=1}^M c_{K-1,j} \phi_{K-1,j}(t)\right)}{1 + \sum_{k=1}^{K-1} \exp\left(\sum_{j=1}^M c_{kj} \phi_{kj}(t)\right)} \\
 w_K &= \frac{1}{1 + \sum_{k=1}^{K-1} \exp\left(\sum_{j=1}^M c_{kj} \phi_{kj}(t)\right)}.
 \end{aligned}$$

Then the partial derivatives of  $w_i, i = 1, \dots, K$  with respect to  $c_{kp}, k = 1, \dots, K-1; p = 1, \dots, M$  are as follows.

$$\begin{aligned}
 \frac{\partial w_1}{\partial c_{kp}} &= -w_1(t)w_k(t)\phi_{kp}(t) + I[k=1]w_k(t)\phi_{kp}(t) \\
 &\quad \vdots \\
 \frac{\partial w_{K-1}}{\partial c_{kp}} &= -w_{K-1}(t)w_k(t)\phi_{kp}(t) + I[k=K-1]w_k(t)\phi_{kp}(t) \\
 \frac{\partial w_K}{\partial c_{kp}} &= -w_K(t)w_k(t)\phi_{kp}(t).
 \end{aligned}$$

And the second partial derivatives of  $w_i, i = 1, \dots, K$  with respect to  $c$  are

$$\begin{aligned}
 \frac{\partial^2 w_1}{\partial c_{kp} \partial c_{lq}} &= w_1(t)(2w_k(t)w_l(t) - I[k=1]w_l(t) - I[l=1]w_k(t) - I[k=l]w_k(t) + I[k=1]I[l=1]) \cdot \\
 &\quad \phi_{kp}(t)\phi_{lq}(t) \\
 &\quad \vdots \\
 \frac{\partial^2 w_{K-1}}{\partial c_{kp} \partial c_{lq}} &= w_{K-1}(t)(2w_k(t)w_l(t) - I[k=K-1]w_l(t) - I[l=K-1]w_k(t) - I[k=l]w_k(t) + \\
 &\quad I[k=K-1]I[l=K-1])\phi_{kp}(t)\phi_{lq}(t) \\
 \frac{\partial^2 w_K}{\partial c_{kp} \partial c_{lq}} &= w_K(t)(2w_k(t)w_l(t) - I[k=l]w_k(t))\phi_{kp}(t)\phi_{lq}(t).
 \end{aligned}$$

The derivatives of  $w_i, i = 1, \dots, K$  with respect to  $t$  are

$$\begin{aligned}\dot{w}_1(t) &= w_1(t) \sum_{j=1}^M c_{1j} \dot{\phi}_{1j}(t) - w_1(t) \left( \sum_{k=1}^{K-1} \left[ w_k(t) \sum_{j=1}^M c_{kj} \dot{\phi}_{kj}(t) \right] \right) \\ &\vdots \\ \dot{w}_{K-1}(t) &= w_{K-1}(t) \sum_{j=1}^M c_{K-1,j} \dot{\phi}_{K-1,j}(t) - w_{K-1}(t) \left( \sum_{k=1}^{K-1} \left[ w_k(t) \sum_{j=1}^M c_{kj} \dot{\phi}_{kj}(t) \right] \right) \\ \dot{w}_K(t) &= -w_K(t) \left( \sum_{k=1}^{K-1} \left[ w_k(t) \sum_{j=1}^M c_{kj} \dot{\phi}_{kj}(t) \right] \right).\end{aligned}$$

And the partial derivatives of  $\dot{w}_i, i = 1, \dots, K$  with respect to  $\mathbf{c}$  are

$$\begin{aligned}\frac{\partial \dot{w}_1}{\partial c_{kp}} &= \frac{\partial w_1}{\partial c_{kp}} \sum_{j=1}^M c_{1j} \dot{\phi}_{1j}(t) + w_1 I[k=1] \dot{\phi}_{1p}(t) - \frac{\partial w_1}{\partial c_{kp}} \left( \sum_{s=1}^{K-1} \left[ w_s \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - w_1 \left( \sum_{s=1}^{K-1} \left[ \frac{\partial w_s}{\partial c_{kp}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - w_1 w_k \dot{\phi}_{kp}(t) \\ &\vdots \\ \frac{\partial \dot{w}_{K-1}}{\partial c_{kp}} &= \frac{\partial w_{K-1}}{\partial c_{kp}} \sum_{j=1}^M c_{K-1,j} \dot{\phi}_{K-1,j}(t) + w_{K-1} I[k=K-1] \dot{\phi}_{K-1,p}(t) - \frac{\partial w_{K-1}}{\partial c_{kp}} \left( \sum_{s=1}^{K-1} \left[ w_s \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - w_{K-1} w_k \dot{\phi}_{kp}(t) \\ \frac{\partial \dot{w}_K}{\partial c_{kp}} &= \frac{\partial w_K}{\partial c_{kp}} \left( \sum_{s=1}^{K-1} \left[ w_s \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - w_K \left( \sum_{s=1}^{K-1} \left[ \frac{\partial w_s}{\partial c_{kp}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - w_K w_k \dot{\phi}_{kp}(t).\end{aligned}$$

Also the second partial derivatives of  $\dot{w}_i, i = 1, \dots, K$  with respect to  $c$  are

$$\begin{aligned}
\frac{\partial^2 \dot{w}_1}{\partial c_{kp} \partial c_{lq}} &= \frac{\partial^2 w_1}{\partial c_{kp} \partial c_{lq}} \sum_{j=1}^M c_{1j} \dot{\phi}_{1j}(t) + \frac{\partial w_1}{\partial c_{lq}} I[k=1] \dot{\phi}_{1p}(t) + \frac{\partial w_1}{\partial c_{kp}} I[l=1] \dot{\phi}_{1q}(t) - \frac{\partial^2 w_1}{\partial c_{kp} \partial c_{lq}} \left( \sum_{s=1}^{K-1} \right. \\
&\quad \left. \left[ w_s \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - \frac{\partial w_1}{c_{lq}} \left( \sum_{s=1}^{K-1} \left[ \frac{\partial w_s}{\partial c_{kp}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - \frac{\partial w_1}{c_{kp}} \left( \sum_{s=1}^{K-1} \left[ \frac{\partial w_s}{\partial c_{lq}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) \\
&\quad - w_1 \left( \sum_{s=1}^{K-1} \left[ \frac{\partial^2 w_s}{\partial c_{kp} \partial c_{lq}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - \frac{\partial w_1}{\partial c_{kp}} w_k \dot{\phi}_{lq}(t) - \frac{\partial w_1}{\partial c_{lq}} w_k \dot{\phi}_{kp}(t) - w_1 \frac{\partial w_k}{\partial c_{lq}} \dot{\phi}_{kp}(t) \\
&\quad - w_1 \frac{\partial w_l}{\partial c_{kp}} \dot{\phi}_{lq}(t) \\
&\vdots \\
\frac{\partial^2 \dot{w}_{K-1}}{\partial c_{kp} \partial c_{lq}} &= \frac{\partial^2 w_{K-1}}{\partial c_{kp} \partial c_{lq}} \sum_{j=1}^M c_{K-1,j} \dot{\phi}_{K-1,j}(t) + \frac{\partial w_{K-1}}{\partial c_{lq}} I[k=K-1] \dot{\phi}_{K-1,p}(t) + \frac{\partial w_{K-1}}{\partial c_{kp}} I[l=K-1] \cdot \\
&\quad \dot{\phi}_{K-1,q}(t) - \frac{\partial^2 w_{K-1}}{\partial c_{kp} \partial c_{lq}} \left( \sum_{s=1}^{K-1} \left[ w_s \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - \frac{\partial w_{K-1}}{c_{lq}} \left( \sum_{s=1}^{K-1} \left[ \frac{\partial w_s}{\partial c_{kp}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - \\
&\quad \frac{\partial w_{K-1}}{c_{kp}} \left( \sum_{s=1}^{K-1} \left[ \frac{\partial w_s}{\partial c_{lq}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - w_{K-1} \left( \sum_{s=1}^{K-1} \left[ \frac{\partial^2 w_s}{\partial c_{kp} \partial c_{lq}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - \frac{\partial w_{K-1}}{\partial c_{kp}} w_k \cdot \\
&\quad \dot{\phi}_{lq}(t) - \frac{\partial w_{K-1}}{\partial c_{lq}} w_k \dot{\phi}_{kp}(t) - w_{K-1} \frac{\partial w_k}{\partial c_{lq}} \dot{\phi}_{kp}(t) - w_{K-1} \frac{\partial w_l}{\partial c_{kp}} \dot{\phi}_{lq}(t) \\
\\
\frac{\partial^2 \dot{w}_K}{\partial c_{kp} \partial c_{lq}} &= - \frac{\partial^2 w_K}{\partial c_{kp} \partial c_{lq}} \left( \sum_{s=1}^{K-1} \left[ w_s \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - \frac{\partial w_K}{c_{lq}} \left( \sum_{s=1}^{K-1} \left[ \frac{\partial w_s}{\partial c_{kp}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - \frac{\partial w_K}{c_{kp}} \left( \sum_{s=1}^{K-1} \right. \\
&\quad \left. \left[ \frac{\partial w_s}{\partial c_{lq}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - w_K \left( \sum_{s=1}^{K-1} \left[ \frac{\partial^2 w_s}{\partial c_{kp} \partial c_{lq}} \sum_{j=1}^M c_{sj} \dot{\phi}_{sj}(t) \right] \right) - \frac{\partial w_K}{\partial c_{kp}} w_k \dot{\phi}_{lq}(t) - \frac{\partial w_K}{\partial c_{lq}} w_k \dot{\phi}_{kp}(t) \\
&\quad - w_K \frac{\partial w_k}{\partial c_{lq}} \dot{\phi}_{kp}(t) - w_K \frac{\partial w_l}{\partial c_{kp}} \dot{\phi}_{lq}(t)
\end{aligned}$$

## APPENDIX C. INNER-OUTER ITERATIVE PROCEDURE IN EM ALGORITHM

In the profiling procedure of estimation for  $(\underline{b}, \underline{c})$  conditioned on  $(\underline{s}^{(t+1)}, \underline{d}^{(t+1)}, \underline{p}^{(t+1)})$  at  $(t+1)$ -th iteration of Chapter 4,  $\lambda$  are being fixed manually and the nuisance parameters are the implicit functions  $\underline{c}$  which are in terms of the structural parameters  $\underline{b}$ . Each time  $\underline{b}$  are changed, an inner function criterion

$$\begin{aligned}
& J(\underline{c}|\underline{b}, \underline{s}^{(t+1)}, \underline{d}^{(t+1)}, \underline{p}^{(t+1)}; \lambda, \underline{\theta}^{\Delta(t)}) \\
&= - \sum_{k_1=1}^K \cdots \sum_{k_n=1}^K p_{1,k_1}^{(t)} \cdots p_{n,k_n}^{(t)} \left\{ (\log p_{k_1}^{(t+1)} + \cdots + \log p_{k_n}^{(t+1)}) + \sum_{i=1}^n \sum_{j=1}^N \sum_{m=1}^M [\log \Gamma(y_{ijm} + \hat{\phi}_{ij}^{-1}) - \log \Gamma(\right. \\
&\quad \left. \hat{\phi}_{ij}^{-1}) - \Gamma(y_{ijm} + 1) - (\hat{\phi}_{ij}^{-1} + y_{ijm}) \log(1 + s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} w_{k_i,j}(\underline{c}) \hat{\phi}_{ij}) + y_{ijm} \log(s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} w_{k_i,j}(\underline{c}) \cdot \right. \\
&\quad \left. \hat{\phi}_{ij})] \right\} + \lambda \sum_{m=1}^{K-1} \int_{[t_1, t_N]} \left[ \frac{dw_m(t|\underline{c})}{dt} - g_m(\underline{w}(t|\underline{c}), \underline{b}) \right]^2 dt
\end{aligned}$$

is optimized with respect to  $\underline{c}$  alone. The estimating function  $\underline{c}$  is regularized by incorporating a penalty term in  $J$  that controls the size of the extent that  $\underline{w}(t|\underline{c}) = \underline{c}'\phi$  fails to satisfy the differential equation exactly. A data fitting criterion

$$\begin{aligned}
& H(\underline{c}(\underline{b}))|\underline{s}^{(t+1)}, \underline{d}^{(t+1)}, \underline{p}^{(t+1)}; \lambda, \underline{\theta}^{\Delta(t)}) \\
&= - \sum_{k_1=1}^K \cdots \sum_{k_n=1}^K p_{1,k_1}^{(t)} \cdots p_{n,k_n}^{(t)} \left\{ (\log p_{k_1}^{(t+1)} + \cdots + \log p_{k_n}^{(t+1)}) + \sum_{i=1}^n \sum_{j=1}^N \sum_{m=1}^M [\log \Gamma(y_{ijm} + \hat{\phi}_{ij}^{-1}) - \log \Gamma(\right. \\
&\quad \left. \hat{\phi}_{ij}^{-1}) - \Gamma(y_{ijm} + 1) - (\hat{\phi}_{ij}^{-1} + y_{ijm}) \log(1 + s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} w_{k_i,j}(\underline{c}) \hat{\phi}_{ij}) + y_{ijm} \log(s_{jm}^{(t+1)} d_{i,k_i}^{(t+1)} w_{k_i,j}(\underline{c}) \cdot \right. \\
&\quad \left. \hat{\phi}_{ij})] \right\}
\end{aligned}$$

is then optimized with respect to the structural parameters  $\underline{b}$  alone. The dependence of  $H$  on  $\underline{b}$  is implicitly through the involvement of  $\underline{c}(\underline{b})$  in defining the fit  $w_i(t|\underline{c})$ . Here we refer to  $J, H$  as inner and outer criteria respectively. In the outer estimation step, the gradient  $\frac{dH}{d\underline{b}}$  is

$$\frac{dH}{d\underline{b}} = \frac{\partial H}{\partial \underline{c}} \frac{d\underline{c}}{d\underline{b}}.$$

Since  $\underline{c}(\underline{b})$  is not available explicitly, we apply the implicit function theorem to obtain

$$\begin{aligned}\frac{d\underline{c}}{d\underline{b}} &= - \left( \frac{\partial^2 J}{\partial^2 \underline{c}^2} \right)^{-1} \frac{\partial^2 J}{\partial \underline{b} \partial \underline{c}}, \\ \frac{dH}{d\underline{b}} &= - \frac{\partial H}{\partial \underline{c}} \left( \frac{\partial^2 J}{\partial^2 \underline{c}^2} \right)^{-1} \frac{\partial^2 J}{\partial \underline{b} \partial \underline{c}}.\end{aligned}$$

The details of the inner-outer iterative procedure is the following:

- (1): Given  $\underline{c}^{(t+1),1} = \underline{c}^{(t)}$  and  $\underline{b}^{(t+1),1} = \underline{b}^{(t)}$ , the optimization of  $J$  with respect to  $\underline{c}$  conditioned on  $\underline{b}^{(t+1),1}$  yields the first estimate  $\underline{c}^{(t+1),2,1}$  of  $\underline{c}$ ,  $\frac{d\underline{c}}{d\underline{b}}$  and  $\frac{dH}{d\underline{b}}$ ;
- (2): Based on  $\underline{b}^{(t+1),1}$  and  $\frac{dH}{d\underline{b}}$  in step (1), use BFGS optimal search formula to derive estimate  $\underline{b}^{(t+1),2,1}$  of  $\underline{b}$ ;
- (3): Repeat steps (1) and (2) by halving BFGS step length until the step size of change in  $\underline{b}$  from BFGS formula is sufficiently small or the criteria  $H$  is not bigger than its value at  $\underline{b}^{(t+1),1}$  of 1st outer iteration (stopping criteria, total B-1 inner iterations). Let the estimates of  $\underline{b}$  and  $\underline{c}$  be  $\underline{b}^{(t+1),2,B}$  and  $\underline{c}^{(t+1),2,B}$  respectively, and denote  $\underline{b}^{(t+1),2,B} = \underline{b}^{(t+1),2}$  and  $\underline{c}^{(t+1),2,B} = \underline{c}^{(t+1),2}$ ;
- (4): Repeat steps (1), (2) and (3) such that either the gradient of  $H$  with respect to  $\underline{b}$  or the change in outer criteria  $H$  is small enough (total M-1 outer iterations). Suppose the estimates obtained are  $\underline{b}^{(t+1),M}$  and  $\underline{c}^{(t+1),M}$ , then denote  $\underline{b}^{(t+1),M} = \underline{b}^{(t+1)}$  and  $\underline{c}^{(t+1),M} = \underline{c}^{(t+1)}$ .

## BIBLIOGRAPHY

- [1] Axelrod, R. and Hamilton, W. (1981). The Evolution of Cooperation. *Science*, *211*, 1390–1396.
- [2] Diggle, S., Griffin, A., Campbell, G. and West, S. (2007). Cooperation and conflict in quorum-sensing bacterial populations. *Nature*, *450*, 411–414.
- [3] Doebeli, M. and Knowlton, N. (1998). The evolution of interspecific mutualisms. *Proc. Natl. Acad. Sci.*, *95*, 8676–8680.
- [4] Durrett, R. and Levin, S. (1997). Allelopathy in Spatially Distributed Populations. *J. theor. Biol.*, *185*, 165–171.
- [5] Gore, J., Youk, H. and van Oudenaarden, A. (2009). Snowdrift game dynamics and facultative cheating in yeast. *Nature*, *459*, 253–256.
- [6] Griffin, A., West, S. and Buckling, A. (2004). Cooperation and competition in pathogenic bacteria. *Nature*, *430*, 1024–1027.
- [7] Hauert, C. (2002). Effect of space in  $2 \times 2$  games. *Int. J. Bifurcation Chaos*, *12*(7), 1531–1548.
- [8] Hauert, C. and Doebeli, M. (2004). Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, *428*, 643–646.
- [9] Huang, Y. and Wu, Z. (2012). Game Dynamic Model for Yeast Development. *Bull. Math. Biol.*, *74*(7), 1469–84.
- [10] Kerr, B., Riley, M., Feldman, M. and Bohannan, B. (2002). Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature*, *418*, 171–174.

- [11] Maclean, R. and Gudelj, I. (2006). Resource competition and social conflict in experimental populations of yeast. *Nature*, 441, 498–501.
- [12] Nowak, M. and May, R. (1992). Evolutionary games and spatial chaos. *Nature*, 359, 826–829.
- [13] Nowak, M. (2006). Five Rules for the Evolution of Cooperation. *Science*, 314, 1560–1563.
- [14] Nowak, M. (2006). *Evolutionary Dynamics*. Cambridge: Belknap Press.
- [15] Rainey, P. and Rainey, K. (2003). Evolution of cooperation and conflict in experimental bacterial populations. *Nature*, 425, 72–74.
- [16] Santorelli, L., Thompson, C., Villegas, E., Svetz, J., Dinh, C., Parikh, A., Sucgang, R., Kuspa, A., Strassmann, J., Queller, D. and Shaulsky, G. (2008). Facultative cheater mutants reveal the genetic complexity of cooperation in social amoebae. *Nature*, 451, 1107–1111.
- [17] Webb, J. (2007). *Game Theory*. London: Springer-Verlag.
- [18] West, S., Griffin, A., Gardner, A. and Diggle, S. (2006). Social evolution theory for microorganisms. *Nature Reviews*, 4, 597–607.
- [19] Abakuks, A. (1980). Conditions for Evolutionarily Stable Strategies. *Journal of Applied Probability*, 17(2), 559–562.
- [20] Bishop, D. T. and Cannings, C. (1976). Models of animal conflict. *Advances in Applied Probability*, 6, 616–621.
- [21] Bishop, D. T. and Cannings, C. (1978). A generalized war of attrition. *Journal of Theoretical Biology*, 70, 85–124.
- [22] Chafee, N. (1975). Asymptotic behavior for solutions of a one-dimensional parabolic equation with homogeneous Neumann boundary conditions. *Journal of Differential equations*, 18, 111–134.
- [23] Ewer, J. P. G. and Peletier, L. A. (1975). On the asymptotic behavior of solutions of semilinear parabolic equations. *Society for Industrial and Applied Mathematics*, 28(1), 43–53.

- [24] Friedman, A. (1964). *Partial differential equations of parabolic type*. Prentice-Hall.
- [25] Garvie, R. M. (2007). Finite-difference schemes for reaction-diffusion equations modeling predator-prey interactions in Matlab. *Bulletin of Mathematical Biology*, 69, 931–956.
- [26] Gore, J., Youk, H. and Oudenaarden van, A. (2009). Snowdrift game dynamics and facultative cheating in yeast. *Nature*, 459, 253–256.
- [27] Hadeler, K. P. (1981). Diffusion in Fisher's poluation model. *Journal of Mathematics*, 11, 39–45.
- [28] Haigh, J. (1975). Game theory and evolution. *Advances in Applied Probability*, 7, 8–11.
- [29] Hofbauer, J. and Sigmund, K. (1998). *Evolutionary games and population dynammics*. Cambridge University Press.
- [30] Hofbauer, J., Schuster, P. and Sigmund, K. (1979). A note on evolutionary stable strategies and game dynamics. *Journal of Theoretical Biology*, 81, 609–612.
- [31] Huang, Y. and Wu, Z. (2012). Game dynamics model for yeast development. *Bulletin of Mathematical Biology*, 74(7), 1469–1484.
- [32] Maynard Smith, J. and Price, G. R. (1973). The logic of animal conflict. *Nature*, 246, 15–18.
- [33] Minura, M. (1979). Asymptotic behavior of a parabolic system related to a planktonic prey and predator model. *Journal of Applied Mathematics*, 37(3), 499–512.
- [34] Nash, J. F. (1951). Non-cooperative games. *Annals of Mathematics*, 54, 286–295.
- [35] Nowak, A. M. (2006). *Evolutionary dynamics*. The Belknap Press of Harvard University Press.
- [36] Pao, V. C. (1999). Numerical analysis of coupled systems of nonlinear parabolic equations. *Journal of Numerical Analysis*, 2, 393–416.
- [37] Redheffer, R., Redlinger, R. and Walter, W. (1988). A theorem of La Salle-Lyapunov type for parabolic systems. *Journal of Mathematical Analysis*, 2, 121–132.

- [38] Taylor, D. P. and Jonker, B. L. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40, 145–156.
- [39] von Neumann, J. and Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton: Princeton University Press.
- [40] Wang M., Huang Y. and Wu Z. (2015). 2D Yeast Cooperation Game. *Submitted*.
- [41] Zeeman, E. C. (1980). Population dynamics from game theory. *Global theory of dynamical systems (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1979)*, Springer-Verlag, 471–497.
- [42] Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.*, 11, R106.
- [43] Bouskill, N. J., Tang, J., Riley, W. J., Brodie, E. L. (2012). Trait-based representation of biological nitrification: Model development testing, and predicted community composition. *Front. Microbiol.*, 3, 364.
- [44] Bullard, H. J., Purdom, E., Hansen, D. K. and Dudoit, S. (2010). Evaluation of statistical methods for normalization and differential expression in mrna-seq experiments. *BMC Bioinformatics*, 11, 94.
- [45] Park B. J. & Lord D. (2008). Adjustment for the maximum likelihood estimate of the negative binomial dispersion parameter. *Texas A & M University*.
- [46] Bullard, H. J., Purdom, E., Hansen, D. K. and Dudoit, S. (2010). Species distribution models: Ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.*, 40, 677-697.
- [47] Faust, K. and Raes, J. (2012). Microbial interactions: From networks to models. *Nat. Rev. Microbiol.*, 10, 538–550.
- [48] Hofbauer, J., Schuster, P. and Sigmund, K. (1979). A Note on Evolutionary Stable Strategies and Game Dynamics. *Journal of Theoretical Biology*, 80, 609–612.

- [49] Hofbauer, J. and Sigmund, K. (1998). *Evolutionary games and population dynamics*. Cambridge University Press, 62–66.
- [50] Jolliffe, I. T. (2002). *Principal Component Analysis*. New York:Springer-Verlag.
- [51] Klitgord, N. and Segre, D. (2010). Environments that induce synthetic microbial ecosystems. *PLOS Comput. Biol.*, 6(11), e1001002.
- [52] Larsen, P. E., Field, D. and Gilbert, J. A. (2010). Predicting bacterial community assemblages using an artificial neural network approach. *Nat. Methods*, 9, 621–625.
- [53] Larowe, D. E., Dale, A. W. and Regnier, P. (2008). A thermodynamic analysis of the anaerobic oxidation of methane in marine sediments. *Geobiology*, 6, 436–449.
- [54] Marino, S., Baxter, N., Huffnagle, G., Petrosino, J. and Schloss, P. (2014). Mathematical modeling of primary succession of murine intestinal microbiota. *PNAS*, 111(1), 439–444.
- [55] Marioni, J. S., Mason, C. E., Mane, S. M., Stephens, M. and Gilad, Y. (2008). Rna-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.*, 18, 1509–1517.
- [56] Miao, H., Xia, X., Perelson, S. A. and Wu, H. (2011). On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics. *SIAM Rev*, 53(1), 3-39.
- [57] Nelder, J. A. and Lee, Y. (1992). Likelihood, quasi-likelihood and pseudolikelihood: some comparisons. *Journal of the Royal Statistical Society, Series B*, 54, 273–284.
- [58] Nowak, M. (2006). Five Rules for the Evolution of Cooperation. *Science*, 314, 1560–1563.
- [59] Ramsay, J., Hooker, G., Campbell, D. and Cao, J. (2007). Parameter estimation for differential equations: a generalized smoothing approach. *J. R. Statist. Soc. B*, 69, 741–796.
- [60] Robinson, M. D., and Smyth, G. K. (2008). Small-sample estimation of negative binomial dispersion, with applications to sage data. *Biostatistics*, 26, 321–332.

- [61] Si, Y., Liu, P., Li, P. and Brutnell, P. T. (2014). Model-based clustering for RNA-seq data. *Bioinformatics*, 30, 197–205.
- [62] Stolyar, S., van Dien, S., Hillesland, K.L., Pinel, N., Lie, T. J., Leigh, J. A., Stahl, D. A. (2007). Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.*, 3, 92.
- [63] Song H. S., Cannon R. W., Beliaev, S. A. and Konopka A. (2014). Mathematical Modeling of Microbial Community Dynamics: A Methodological Review. *Processes*, 2(4), 711–752.
- [64] Taylor, D. P. and Jonker, B. L. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40, 145–156.
- [65] Zeeman, E. C. (1980). *Population dynamics from game theory*, in *Global theory of dynamical systems*. Springer Lecture Notes in Mathematics Vol. 819. Heidelberg: Springer.
- [66] Besag, J. (1974). Spatial interaction and the analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, 36, 192–236.
- [67] Brown. J. S., Sanderson, M. J., and Michod, R. E. (1982). Evolution of social behavior by reciprocation. *Journal of theoretical biology*, 99, 319–339.
- [68] Cressie, N. (1993). *Statistics for spatial data*. Wiley Interscience.
- [69] Durrett, R. and Levin, S. (1997). Allelopathy in spatially distributed populations. *J. Theor. Biol.*, 185, 165–171.
- [70] Ferguson, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press.
- [71] Ferriere, R. and Michod, R. E. (2000). *Wave patterns in spatial games and the evolution of cooperation*. The geometry of ecological interaction, simplifying spatial complexity, Eds. Dieckmann, U., Law, R., and Metz, J.A.J., Cambridge.
- [72] Gaetan, C. and Guyon, X. (2010). *Spatial statistics and modeling*. Springer.

- [73] Hauert, C. H. (2002). Effects of space in  $2 \times 2$  games. *International Journal of Bifurcation and Chaos*, 12(7), 1531–1548.
- [74] Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge.
- [75] Kaiser, M. S. and Cressie, N. (2000). The construction of multivariate distributions from Markov random fields. *Journal of Multivariate Analysis*, 73, 199–220.
- [76] Kaiser, M. S. and Caragea, P. (2007). *Exploring dependence with data on spatial lattice*. Technical report, Iowa State University.
- [77] Kaiser, M. S. and Caragea, P. (2009). Exploring dependence with data on spatial lattice. *Biometrics*, 65, 857–865.
- [78] Kaiser, M. S., Lahiri, S. N. and Nordman, D. J. (2012). Goodness of fit tests for a class of Markov random field models. *Annals of Statistics*, 40(1), 104–130.
- [79] Kaiser, M. S., Pazdernik, K. T., Hoeksema, A. B., and Nutter, F. W. (2014). Modeling the spread of plant disease using a sequence of binary random fields with absorbing state. *Spatial Statistics*, 9, 38–50.
- [80] Nowak, M. A. and May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, 359, 826–829.
- [81] Nowak, M. A., Bonhoeffer, S. and May, R. M. (1994). Spatial games and maintenance of cooperation. *P.Natl.Acad.Sci.USA.*, 91, 4877–4881.
- [82] Nowak, M. A. (2008). *Evolutionary dynamics: exploring the equations of life*. Belknap Press of Harvard University Press.
- [83] Roca, C. P., Cuesta, J. A. and Sanchez, A. (2009). Effect of spatial structure on the evolution of cooperation. *Physical review E*, 80, 046106.
- [84] Berman, A. and Plemmons, R. J. (1979). *Nonnegative Matrices in the Mathematical Sciences*. Academic Press.

- [85] Bomze, I. (2002). Regularity vs. degeneracy in dynamics, games, and optimization: A unified approach to different aspects. *SIAM Review*, 44, 394–414.
- [86] Borwein, J. M. (1982). Necessary and sufficient conditions for quadratic minimality. *Numer. Funct. Anal. Optim.*, 5, 127–140.
- [87] Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.
- [88] Campbell, N. A., Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V. and Jackson, R. B. (2008). *Biology, 8th edition*. Pearson Education Inc..
- [89] Ewens, W. J. (2004). *Mathematical Population Genetics*. New York: Springer-Verlag.
- [90] Fisher, R. A. (1930). *The Genetic Theory of Natural Selection*. Oxford: Clarendon Press.
- [91] Hardy, G. H. (1908). Mendelian proportions in a mixed population. *Zeitschrift fur Induktive Abstammungs- und Vererbungslehre*, 1, p395.
- [92] Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.
- [93] Maynard Smith, J. and Price, G. R. (1973). The logic of animal conflict. *Nature*, 246, 15–18.
- [94] Motzkin, T. and Straus, E. (1965). Maxima for graphs and a new proof of a theorem of Turán. *Canadian J. Math.*, 17 (1965), 533–540.
- [95] Murty, K. G. and Kabadi, S. N. (1987). Some NP-complete problems in quadratic and linear programming. *Math. Programming*, 39 (1987), 117–129.
- [96] Nash, J. (1950). Equilibrium points in  $n$ -person games. *Proceedings of the National Academy of Sciences*, 36, 48–49.
- [97] Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. New York: Springer-Verlag.
- [98] Pardalos, P., Ye, Y. and Han, C. (1991). Algorithms for the solution of quadratic knapsack problems. *Linear Algebra and Its Applications*, 152, 69–91.

- [99] Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. The MIT Press.
- [100] Templeton, A. R. (2006). *Population Genetics and Microevolutionary Theory*. John Wiley & Sons Inc..
- [101] Trefethen, L. N. and Bau III, D. (1997). *Numerical Linear Algebra*. SIAM, 1997.
- [102] Weibull, J. W. (1995). *Evolutionary Game Theory*. The MIT Press.
- [103] Weinberg, W. (1908). Aber den nachweis der vererbung beim menschen. *Jahreshefte des Vereins für vaterlandische Naturkunde in Wurttemberg*, 64, 368–382.
- [104] Bomze, I. M., Budinich, M., Pardalos, P. M., and Pelillo, M. (1999). The maximum clique problem. *Handbook of Combinatorial Optimization 4*, Kluwer Academic Publishers, 1–74.
- [105] Bomze, I. (2002). Regularity vs. degeneracy in dynamics, games, and optimization: A unified approach to different aspects. *SIAM Review*, 44, 394–414.
- [106] Broom, M., Cannings, C., and Vickers, G. T. (1993). On the number of local maxima of a constrained quadratic form. *Proceedings of the Royal Society of London, Series A*, 443, 573–584.
- [107] Fletcher, R. (2000). *Practical Methods of Optimization*. Wiley.
- [108] Garey, M. R. and Johnson, D. S. (1979). *Computer and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company.
- [109] Hidalgo, C. (2011). *Networks, Complexity and Its Applications*. Springer.
- [110] Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.
- [111] Huang, Y., Hao, Y., Wang, M., Zhou, W., and Wu, Z. (2015). Optimality and stability of symmetric evolutionary games with applications in genetic selection. *Journal of Mathematical Biosciences and Engineering*, 12(3).
- [112] Lewis, T. G. (2009), *Network Science: Theory and Applications*. Wiley.

- [113] Motzkin, T. S. and Straus, E. G. (1965). Maxima for graphs and a new proof of a theorem of Turan. *Canadian Journal of Mathematics*, 17, 533–540.
- [114] Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer.
- [115] Pardalos, P., Ye, Y., and Han, C. (1991). Algorithms for the solution of quadratic knapsack problems. *Linear Algebra and Its Applications*, 152, 69–91.
- [116] Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. The MIT Press.
- [117] Scott, J. (2000). *Social Network Analysis: A Handbook*. SAGE Publication Ltd.
- [118] Vickers, G. T. and Cannings, C. (1988). On the number of stable equilibria in a one-locus, multi-allelic system. *Journal of Theoretical Biology*, 131, 273–277.
- [119] Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- [120] Weibull, J. W. (1995). *Evolutionary Game Theory*. The MIT Press.

## PUBLICATION LIST

### Peer Reviewed Journal Articles

- [1] Y. Huang, Y. Hao, **M. Wang**, W. Zhou and Z. Wu. Optimality and stability of symmetric evolutionary games with applications in genetic selection. *Mathematical Biosciences and Engineering*, 12(3).
- [2] **M. Wang**, Y. Huang, Y. Hao, W. Zhou and Z. Wu. Evolution of social cliques. *Journal of Optimization Theory and Applications*, *under review*.
- [3] **M. Wang**, Y. Huang and Z. Wu. 2D Yeast cooperation games model. *To be submitted*.
- [4] **M. Wang** and Z. Wu, ESS in Nonlinear Game and the Adapted Replicator Diffusion Equation. *To be submitted*.

### Papers in Preparation

- [5] **M. Wang** and K. Dorman. Parameter estimation in dynamical metagenomic model. *In preparation*.
- [6] **M. Wang**, W. Zhou, et al. A statistical method on modeling of evolutionary games: Markov random field model. *In preparation*.