Title
Mathematical Modeling of Cooperation Based Diversification and Speciation

Author
Wood, Karen Elizabeth

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Mathematical Modeling of Cooperation Based Diversification and Speciation

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Mathematics

by

Karen Elizabeth Wood

Dissertation Committee:
Professor Natalia L. Komarova, Chair
Chancellor's Professor Hongkai Zhao
Associate Professor German Enciso Ruiz

2017

DEDICATION

To my grandmas, Jean and Nancy, both of whom always supported my continued education and have encouraged my interest in mathematics over the years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my committee chair and advisor, Professor Natalia L. Komarova. Throughout my time at UC Irvine, she has been a constant source of support and guidance for subjects both within academia and without.

I would also like to thank the other members of my thesis committee, Professor German A. Enciso Ruiz and Professor Hongkai Zhao, for their support as mentors and as educators during my time at UCI.

Financial support was provided by Graduate Assistance in Areas of National Need (GAANN) and the Diverse Educational Community and Doctoral Experience (DECADE) at UC Irvine.

I am humbled by the support that I have received by my parents, family and friends. Many helped me by proof reading papers, others by providing emotional support during rough times, and all were incredibly patient with me over the years as I pursued my dream. I would especially like to thank my husband, Craig White, who has been incredibly supportive throughout this crazy ride. Thank you.

CURRICULUM VITAE

Karen Elizabeth Wood

EDUCATION

Doctor of Philosophy in Mathematics **2017**
University of California, Irvine *Irvine, California*

Master of Science in Mathematics **2013**
University of California, Irvine *Irvine, California*

Bachelor of Science in Mathematics **2010**
California Polytechnic University, Pomona *Pomona, California*

RESEARCH EXPERIENCE

Graduate Research Assistant **2013–2017**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2011–2016**
University of California, Irvine *Irvine, California*

ABSTRACT OF THE DISSERTATION

Mathematical Modeling of Cooperation Based Diversification and Speciation

By

Karen Elizabeth Wood

Doctor of Philosophy in Mathematics

University of California, Irvine, 2017

Professor Natalia L. Komarova, Chair

Cooperation in biology and diversification of species have been widely studied by both evolutionary biologists and mathematicians. In this work we examine both of these seemingly unrelated phenomena and propose that there could be a context where they are connected. We focus on a setting where individuals in a shared environment cooperate by sharing products of two distinct parts of a complex task. Different strategies can evolve: individuals can complete all parts of the complex task, choosing self-sufficiency over cooperation, or they may choose to split parts of the task and share the products for mutual benefit, such that distinct groups of the organisms specialize on a subset of elementary tasks. We first examine this possibility using a quasispecies system, and then by using the methodology of adaptive dynamics, both analytically and by stochastic agent-based simulations, to investigate the conditions where branching into distinct cooperating subgroups occurs. We show that if performing multiple tasks is associated with additional cost, branching occurs for a wide parameter range, and is stable against the invasion of non-cooperating non-producers ("cheaters"). We hypothesize that over time, this can lead to evolutionary speciation, providing a novel mechanism of speciation based on cooperation. In addition, we investigate whether microscopic assumptions of the interaction rules of the simulations may play a role in the resulting dynamics. To do this, we derive ordinary differential equations for the mean trait values for four models, which differ by (1) the number of interactions each individual

engages in before the payoff is determined (interacting with the entire population vs interacting with one randomly chosen individual), and (2) the type of criterion (probabilistic vs deterministic) by which the winner of each competition is determined. We find that the mean trait dynamics are the same in all four cases when only one trait in a population of cooperators is evolving. However, when we include "cheaters" we find, surprisingly, that the rules do make a difference, and the steady state to which the system converges can depend both on the number of interactions and on the criterion for determining the winners.

Chapter 1

Introduction

This work will investigate the evolutionary dynamics of populations in the presence of cooperation within a competitive environment. The fascinating topic of cooperation in biology has been studied in depth by others, as it is both puzzling, and has wide ranging implications. Biologists have examined cooperation in many circumstances, including division of labor in cancerous tumors [1], within biofilms [52, 5], within bacteria populations [26], and within yeast populations [46], to give just a few examples.

We were first inspired by cooperation of cells observed in some cancers [1]. These cancer cells were able to specialize, performing a subset of tasks associated with cancer, rather than performing all the tasks that a cancerous population performs. With this in mind, we developed a system of differential equations to model cooperation and competition. We assumed the existence of two separate tasks, whose products increase individuals' fitness. In our model we considered several sub-populations, each characterized by a different behavior, identified by how many tasks are performed: one, both, or none of the tasks. The model captures the changes in the fraction of the population of each behavior, based on the presence of the other cooperative populations. Using this concept, we were able to analyze the steady states of the system across a wide variety of assumptions, including varying parameters and the functional forms characterizing the cooperation within the model. We also introduced mutations between the different populations and examined the effects on the population dynamics. We found that depending on parameters, different equilibria were stable. Some were characterized by the generalists (species that could perform both tasks) taking over, and others by the coexistence of specialists (two species that each were able to perform only one of the two tasks, but were willing to share the products, thus maintaining stable coexistence). It was also possible that a generalist could coexist with one of the specialists; furthermore, for some parameter sets, bistability between the generalist-only equilibrium and the other coexisting mixed populations was possible. The interesting result, which is the coexistence equilibrium, was observed over a large parameter range. This suggests that the evolutionary strategy to specialize on one of the task is enabled by the existence of cooperation among individuals, which brought us to the next question: is it possible that cooperation can be responsible for diversification / speciation in evolutionary dynamics?

In order to study this question, we had to make the model more flexible. In the first model, each population has a fixed behavior that does not change over time. Now, to study the evolutionary pathway to speciation, we needed to include the possibility of a gradual change of traits. In particular, we considered two evolving traits which represented the amount of product that an individual produced (and shared) for each of the two sub-tasks. In other words, an individual's investment into each task was quantified by a trait value, with each individual being represented by two trait values. These individuals die, reproduce, and compete all within in a shared environment. Their reproductive success is defined by the payoff they receive by utilizing the products of the two tasks (either produced by the individual or obtained from another individual by sharing, as in the division of labor game). More precisely, individuals are involved in a number of interactions with others in the population, and their payoff is calculated based on the results of these interactions. This is measured based on the amount of products received, which is governed by a fitness function.

Instead of focusing on a particular functional for of this fitness function, we identified general biologically meaningful requirements that this function should satisfy. For example, it was assumed that the benefit was received only if enough of both products were received; further, the existence of two types of cost were postulated: a cost associated with performing either of the tasks, and an additional cost resulting from being able to participate in both tasks, which is only incurred if a sufficient amount of both products is being made. This second type of cost reflects additional costs for maintaining the ability to perform two behaviors, or costs associated with the effort required to switch behaviors. The adaptive dynamics of the resulting system was investigated both analytically and numerically, by using specific functional forms and performing stochastic agent-based simulations. We found that while some parameter regimes lead to the individuals performing both tasks, others lead to the population branching into multiple sub-populations, each specializing in one of the tasks. While examining this model, we also investigated how cheaters, which can benefit from the behavior of non-cheaters, affect the system. In this context, cheaters are non-producing, non-cooperators. In some cases, cheaters drove the non-cheaters to extinction, but in other cases, under surprisingly wide parameter regimes, cheaters were able to coexist with non-cheaters, and the system with cheaters exhibited the same cooperation driven branching. This showed that the mechanism of speciation based on cooperation is robust with respect to the intrusion of cheaters.

While working on the different scenarios described above, we discovered numerically that although the speciation phenomenon was very robust, some other aspects of evolutionary behavior (such as post-branching behavior) were affected by small and seemingly insignificant details of the model formulation. To explain this, we note that in models where the fitness of individuals is calculated based on interactions with others, one can make different assumptions on how many interactions happen per individual. On the one hand, an infinite number of interactions (or as many as there are agents in the population) is often assumed in the literature. On the other hand, in simulations, a finite (and often small) number of interactions per individual is often taken into account. It is tacitly assumed that the number of interactions does not change the behavior. In our work, we tried to vary the number of interactions and found that there were differences. This inspired us to look further into this phenomenon. To study it systematically, we focused on four types of models, which differed by (1) the number of interactions each individual engages in before the payoff is determined (interacting with the entire population vs interacting with one randomly chosen individual), and (2) the criterion by which the winner of each competition is determined: the deterministic criterion, where the individual with a larger payoff wins, vs a probabilistic criterion, where the probability to win is proportional to the payoff difference. In order to understand the influence of these choices on the systems' behavior, we derived, from first principles, the ODEs describing the mean trait evolution. It turned out that in the system of cooperators where only one trait was evolving, all four models were characterized by the same ODE, and thus we concluded that choices (1) and (2) did not make a difference on the behavior. Interestingly, once we slightly increased the complexity of the system by adding a fraction cheaters to the population, these choices began to matter. In particular, we found that the equilibrium solution depended on assumptions (1) and (2). This counterintuitive result can be used by future modelers, alerting them to the possibility that insignificant modeling choices might matter when studying evolutionary dynamics.

Chapter 2

Coexistence of Cooperators with Different Levels of Investment

2.1 Introduction

It has been asserted that for a cell to become cancer it must first obtain the "hallmarks" of cancer [23]. As cells obtain these mutations they can be grouped into subpopulations identified by the collection of "hallmarks" they've collected. These subpopulations may be cooperating with one another by producing a shared resource; for example, they may produce different chemicals or substances that another subpopulation needs and by coexisting together they are able to survive and proliferate [1]. We chose to examine this behavior through the use of a system of differential equations which models interaction between several populations.

2.2 Model

To model cooperation and competition, we formulated system (2.1) which has the form of a quasispecies model (see appendix A). We consider a system with four populations: ab, aB, Ab and AB, with base growth rates r_{ab}, r_{aB}, r_{Ab} and r_{AB}, respectively. The system includes mutation from type ab to aB or Ab (types aB and Ab will be henceforth referred to as once-mutated), and from type aB or Ab to type AB (twice-mutated). In addition, three of the population types: aB, Ab, and AB, cooperate through the functions g_{aB} and g_{Ab} which increase growth rates in the presence of other populations. Specifically, population types aB and AB improve the growth rate of population type Ab through function $g_{aB}(x_{Ab}, x_{AB})$ and the population types Ab and AB help the growth rate of population type aB through function $g_{Ab}(x_{aB}, x_{AB})$. The cell type ab could be considered a healthy cell, aB and Ab as cells that do not yet have all the hallmarks of cancer, but can work together to represent the full range of hallmarks, and type AB is a cell type with all the hallmarks of cancer. These interactions are summarized in figure 2.1.

$$\dot{x_{ab}} = r_{ab}x_{ab}(1 - u_1 - u_2) - \phi x_{ab} \tag{2.1a}$$

$$\dot{x_{Ab}} = [r_{Ab} + g_{Ab}(x_{aB}, x_{AB})]x_{Ab}(1 - u_2) + r_{ab}x_{ab}u_1 - \phi x_{Ab} \tag{2.1b}$$

$$\dot{x_{aB}} = [r_{aB} + g_{aB}(x_{Ab}, x_{AB})]x_{aB}(1 - u_1) + r_{ab}x_{ab}u_2 - \phi x_{aB} \tag{2.1c}$$

$$\dot{x_{AB}} = r_{AB}x_{AB} + [r_{Ab} + g_{Ab}(x_{aB}, x_{AB})]x_{Ab}u_2 + [r_{aB} + g_{aB}(x_{Ab}, x_{AB})]x_{aB}u_1 - \phi x_{AB} \tag{2.1d}$$

The growth rate of population types aB and Ab when we include the cooperation functions, are $r_{aB} + g_{aB}(x_{Ab}, x_{AB})$ and $r_{Ab} + g_{Ab}(x_{aB}, x_{AB})$, so the average of the growth rates of the population is

$$\phi = r_{ab}x_{ab} + [r_{aB} + g_{aB}(x_{Ab}, x_{AB})]x_{aB} + [r_{Ab} + g_{Ab}(x_{aB}, x_{AB})]x_{Ab} + r_{AB}x_{AB}.$$

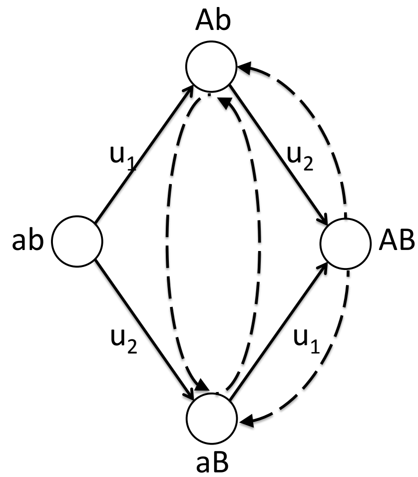


Figure 2.1: The schematic of mutations and cooperation. Mutation edges are solid arrows, with mutation rates indicated. Dashed arrows indicate the directions in which cooperation occurs through the functions g_{Ab} and g_{aB}.

As this model is very complex, we would like to begin by examining the following simplified versions.

2.2.1 Asymmetric Cooperation

To study the role of cooperation in system heterogeneity, let us first consider the case in which cooperation is asymmetric, as in figure 2.2. We will make a few assumptions about the growth rates to examine the behavior that we are interested in. In particular, we will assume that the base growth rates of populations of type aB and Ab are not larger than that of ab. These restrictions are inspired by the idea that effort may be required to produce products A and B. For this asymmetric case, the presence of populations which produce product B (populations AB and aB) improves the growth of population Ab. The population type AB produces both A and B, so it does not have to interact with others for an enhanced growth rate. A reasonable growth rate of the population type AB could be a growth rate similar to that of population Ab in the presence of populations which produce product B.

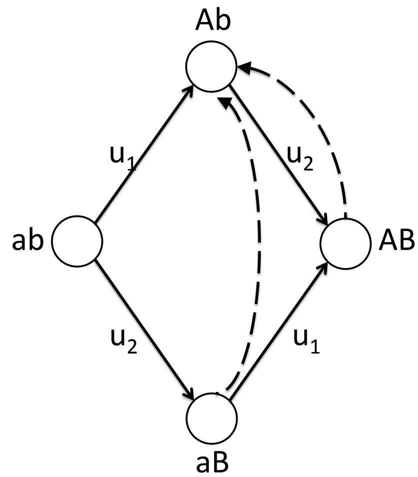


Figure 2.2: The schematic of mutations and cooperation for the system with asymmetric cooperation. Mutation edges are solid arrows, with mutation rates indicated. Dashed arrows indicate the directions in which cooperation occurs through the function g_{Ab}.

The system of ODEs that describes these processes is the following:

$$\dot{x_{ab}} = r_{ab}(1 - u_1 - u_2)x_{ab} - \phi x_{ab} \tag{2.2a}$$

$$\dot{x_{aB}} = r_{ab}u_2 x_{ab} + r_{aB}(1 - u_1)x_{aB} - \phi x_{aB} \tag{2.2b}$$

$$\dot{x_{Ab}} = r_{ab}u_1 x_{ab} + [r_{Ab} + g_{Ab}(x_{aB}, x_{AB})](1 - u_2)x_{Ab} - \phi x_{Ab} \tag{2.2c}$$

$$\dot{x_{AB}} = [r_{Ab} + g_{Ab}(x_{aB}, x_{AB})]u_2 x_{Ab} + r_{aB}u_1 x_{aB} + r_{AB}x_{AB} - \phi x_{AB} \tag{2.2d}$$

For this system, $g_{Ab}(x_{aB}, x_{AB})$ is a non-decreasing function. The function ϕ is defined as before, it is based on the growth rates and cooperation functions. Under the assumptions that $r_{Ab} + g_{Ab}(x_{aB}, x_{AB})$ is greater than both r_{aB} and r_{ab}, and r_{AB} is greater than r_{Ab} and r_{aB}, the system will be dominated by types Ab and/or AB since they will have the highest growth rate. This simple analysis does not provide the details we are interested in, such as if Ab and AB can coexist and under which circumstances this could occur. To do so, we will remove the mutations and consider a linear form for the cooperation function. By the work done in appendix A, we know that so long as mutation rates are low, the system without mutations will have steady states similar to those when there are mutations.

$(x_{ab}, x_{aB}, x_{Ab}, x_{AB})$	Stability Conditions
$(1, 0, 0, 0)$	$r_{aB} < r_{ab},\ r_{Ab} < r_{ab},\ r_{AB} < r_{ab}$
$(0, 0, 0, 1)$	$r_{ab} < r_{AB},\ r_{aB} < r_{AB},\ r_{Ab} + \alpha < r_{AB}$
$(0, 1, 0, 0)$	$r_{ab} < r_{aB},\ r_{Ab} + \alpha < r_{aB},\ r_{AB} < r_{aB}$
$(0, 0, 1, 0)$	$r_{ab} < r_{Ab},\ r_{aB} < r_{Ab},\ r_{AB} < r_{Ab}$
$(0, \frac{r_{aB}-r_{Ab}}{\alpha}, \frac{\alpha - r_{aB} + r_{Ab}}{\alpha}, 0)$	$r_{ab} < r_{AB},\ r_{AB} < r_{aB}$ and positive populations
$(0, 0, \frac{r_{AB}-r_{Ab}}{\alpha}, \frac{\alpha + r_{Ab} - r_{AB}}{\alpha})$	$r_{ab} < r_{AB},\ r_{aB} < r_{AB}$ and positive populations

Table 2.1: Steady states and their stability conditions for the system 2.3.

2.2.1.1 No Mutations, Linear Function for Cooperation Function

The next simplified model we will consider has two further simplifications: no mutations can occur, and the cooperation function is linear, $g_{Ab}(x_{aB}, x_{AB}) = \alpha(x_{aB} + x_{AB})$. The system is now:

$$\dot{x_{ab}} = r_{ab}x_{ab} - \phi x_{ab} \tag{2.3a}$$

$$\dot{x_{aB}} = r_{aB}x_{aB} - \phi x_{aB} \tag{2.3b}$$

$$\dot{x_{Ab}} = [r_{Ab} + \alpha(x_{aB} + x_{AB})]x_{Ab} - \phi x_{Ab} \tag{2.3c}$$

$$\dot{x_{AB}} = r_{AB}x_{AB} - \phi x_{AB} \tag{2.3d}$$

The function ϕ is defined as before, with $\phi = r_{ab}x_{ab} + r_{aB}x_{aB} + [r_{Ab} + \alpha(x_{aB} + x_{AB})]x_{Ab} + r_{AB}x_{AB}$. The steady states of system (2.3 can be found by setting the differential equations equal to zero. By examining the eigenvalues of the Jacobian of the system at these steady states, we can determine the stability conditions. The exploration of these eigenvalues is contained in appendix A. The steady states and their stability conditions are listed in table 2.1.

Steady states of the forms $(0, 1, 0, 0)$, $(0, 0, 1, 0)$, and $(0, x_{aB}, x_{Ab}, 0)$ are not stable under our assumptions. Specifically, the steady state of the form $(0, 1, 0, 0)$ requires $r_{aB} > r_{AB}$ and $r_{aB} > r_{AB}$, but in the model, we would like cell type aB to have a smaller base growth rate than the twice-mutated cell type AB. Similar reasoning applies to states $(0,0,1,0)$ and $(0, x_{aB}, x_{Ab}, 0)$. By considering the restrictions on the relationships between the growth rates, and assuming positive parameters, growth rates and populations, we have that only $(1,0,0,0)$, $(0,0,0,1)$, and $(0,0,x_{Ab},x_{AB})$ are stable steady states. When $r_{ab} > r_{AB}$, the steady state $(1,0,0,0)$ is the only stable one. If instead $r_{AB} > r_{ab}$, we have that $(0,0,0,1)$ is stable, and if additionally $r_{AB} > \alpha + r_{Ab}$, then $(0,0,x_{Ab},x_{AB})$ may also be stable. Thus, when $r_{AB} > r_{ab}$, there may be bistability, and the stable steady state will depend on the initial conditions.

2.2.2 Symmetric Cooperation

We will now consider symmetric cooperation, the same as shown in figure 2.1. We will start with first examining simplified cases.

2.2.2.1 Linear Cooperation Function, Three Populations

We will examine the case with only three populations. In particular, we will consider the case without any of population type AB. This would be equivalent to assuming that $x_{AB} = 0$ in system 2.1, and corresponds to figure 2.3:

$$\dot{x_{ab}} = r_{ab}x_{ab} - \phi x_{ab} \tag{2.4a}$$

$$\dot{x_{Ab}} = r_{Ab}x_{Ab} + \alpha x_{aB}x_{Ab} - \phi x_{Ab} \tag{2.4b}$$

$$\dot{x_{aB}} = r_{aB}x_{aB} + \alpha x_{Ab}x_{aB} - \phi x_{aB} \tag{2.4c}$$

Similarly to before, we have $\phi = r_{ab}x_{ab} + (r_{Ab} + \alpha x_{aB})x_{Ab} + (r_{aB} + \alpha x_{Ab})x_{aB}$.

We will consider r_{ab} larger than both r_{Ab} and r_{aB}, as we would like to examine the case

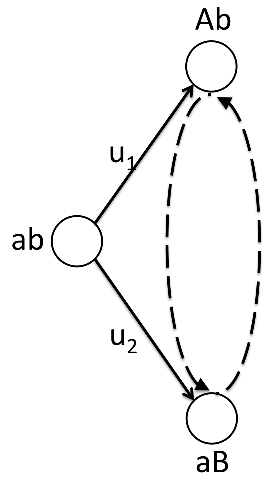


Figure 2.3: The schematic of mutations and cooperation for the system with symmetric cooperation, but only 3 populations. Mutation edges are solid arrows, with mutation rates u_1 and u_2 in the directions indicated. Dashed arrows indicate the directions in which cooperation occurs through the functions g_{Ab} and g_{aB}.

$(x_{ab}, x_{aB}, x_{Ab}, x_{AB})$	Stability Conditions
$(1, 0, 0)$	$r_{Ab} < r_{ab},\ r_{aB} < r_{ab}$
$(0, 1, 0)$	$r_{ab} < r_{Ab},\ r_{aB} + \alpha < r_{Ab}$
$(0, 0, 1)$	$r_{ab} < r_{aB},\ r_{Ab} + \alpha < r_{aB}$
$(\frac{\alpha-2r_{ab}+r_{Ab}+r_{aB}}{\alpha}, \frac{r_{ab}-r_{aB}}{\alpha}, \frac{r_{ab}-r_{Ab}}{\alpha})$	unstable if populations are positive
$(0, \frac{\alpha+r_{Ab}-r_{aB}}{2\alpha}, \frac{\alpha-r_{Ab}+r_{aB}}{2\alpha})$	$2r_{ab} < \alpha + r_{Ab} + r_{aB}$ and positive populations

Table 2.2: Steady states and their stability conditions for the system 2.4

where the cell types Ab and aB do not perform as well as cell type ab in the absence of the paired assisting cell as discussed before.

The steady states of the system defined in system 2.4 can be found and we can determine the stability conditions, as done in appendix A. The steady states and their stability conditions are listed in 2.2.

The states (0,1,0) and (0,0,1) are not stable under the assumptions we have made about the growth rates. There is overlap between the conditions for the stability of the steady states (1,0,0) and $(0, x_{Ab}, x_{aB})$, such that if $\alpha > 2r_{ab} - r_{Ab} - r_{aB}$, both steady states have their stability conditions met. This means that for some parameter regimes, we have bistability.

2.2.2.2 Linear Functions for Cooperation Functions, Four Populations

We will now consider the case with four populations. In this system, the once-mutated types, x_{aB} and x_{Ab}, are both helped by x_{AB} and x_{Ab} or x_{aB}, respectively. We will still consider a linear cooperation function. The system is now in the form of system (2.1), without mutation.

$$\dot{x_{ab}} = r_{ab}x_{ab} - \phi x_{ab} \tag{2.5a}$$

$$\dot{x_{aB}} = [r_{aB} + \beta(x_{Ab} + x_{AB})]x_{aB} - \phi x_{aB} \tag{2.5b}$$

$$\dot{x_{Ab}} = [r_{Ab} + \alpha(x_{aB} + x_{AB})]x_{Ab} - \phi x_{Ab} \tag{2.5c}$$

$$\dot{x_{AB}} = r_{AB}x_{AB} - \phi x_{AB} \tag{2.5d}$$

In this system, we are assuming a linear form for the functions g_{aB} and g_{Ab}: $g_{aB}(x_{Ab}, x_{AB}) = \beta(x_{Ab} + x_{AB})$ and $g_{Ab}(x_{aB}, x_{AB}) = \alpha(x_{aB} + x_{AB})$ as in section 2.2.1.1. We are still making the same assumptions about the relations of the growth rates, so, $r_{ab} > r_{aB}$, $r_{ab} > r_{Ab}$, $r_{AB} > r_{aB}$, and $r_{AB} > r_{Ab}$. The steady states and the conditions for stability were found, and are listed in table 2.3. For ease of representation of the steady states, we will use $w = r_{AB} - r_{Ab}$ and $z = r_{AB} - r_{aB}$, so $r_{aB} - r_{Ab} = w - z$. Again, the eigenvalues are listed in appendix A.

The steady states where one of the single mutated cells becomes the only surviving populations, (0,1,0,0) and (0,0,1,0) are not stable under our assumptions. We would have to loosen the requirements to allow both $r_{aB} > r_{ab}$ and $r_{aB} > r_{AB}$ for steady state (0,1,0,0) or $r_{Ab} > r_{ab}$ and $r_{Ab} > r_{AB}$ for steady state (0,0,1,0) to become stable.

The relationship between the parameters and the stable fixed points can be more clearly expressed using a graph as shown in figure 2.4. The point which is stable is dependent on $w = r_{AB} - r_{Ab}$ and $z = r_{AB} - r_{aB}$ and is indicated on its region on the graph. The graph was

$(x_{ab}, x_{aB}, x_{Ab}, x_{AB})$	Stability Conditions
$(1,0,0,0)$	Stable if $r_{ab} > r_{AB}$
$(0,1,0,0)$	Unstable under our assumptions
$(0,0,1,0)$	Unstable under our assumptions
$(0,0,0,1)$	$r_{ab} < r_{AB},\ \beta < z,\ \alpha < w$
$(0, \frac{\beta+w-z}{\alpha+\beta}, \frac{\alpha+z-w}{\alpha+\beta}, 0)$	$r_{AB} < \frac{\alpha(\beta+r_{aB})+\beta r_{Ab}}{\alpha+\beta},\ \beta > \frac{\alpha(r_{aB}-r_{ab})}{r_{ab}-r_{Ab}-\alpha}$
$(0, \frac{\beta-z}{\beta}, 0, \frac{z}{\beta})$	$r_{ab} < r_{AB}$, $\alpha < w$ and positive populations
$(0, 0, \frac{\alpha-w}{\alpha}, \frac{w}{\alpha})$	$r_{ab} < r_{AB},\ \beta < z$ and positive populations
$(0, \frac{\beta-z}{\beta}, \frac{\alpha-w}{\alpha}, \frac{-\alpha(\beta-z)+\beta w}{\alpha\beta})$	$r_{ab} < r_{AB}$ and positive populations
$(\frac{\alpha(\beta-r_{ab}+r_{aB})+\beta(-r_{ab}+r_{Ab})}{\alpha\beta}, \frac{r_{ab}-r_{Ab}}{\alpha}, \frac{r_{ab}-r_{aB}}{\beta}, 0)$	Unstable if populations are positive

Table 2.3: Steady states and their stability conditions for the system 2.5

created using the parameters $\alpha = .6$ and $\beta = .5$, and the line separating the steady states $(0, x_{aB}, x_{Ab}, 0)$ and $(0, x_{aB}, x_{Ab}, x_{AB})$ is $z = -\frac{6}{5}w + .6$.

When $r_{AB} < r_{ab}$, the fixed point $(1,0,0,0)$ is always stable since under our assumptions, the once-mutated cells will have a smaller growth rate than the unmutated cells. In addition, the steady state $(0, x_{aB}, x_{Ab}, 0)$ is bistable with $(1,0,0,0)$ when $\frac{\alpha(\beta+r_{aB})+\beta r_{Ab}}{\alpha+\beta} > r_{ab}$.

The differences between r_{AB} and r_{aB} and between r_{AB} and r_{Ab} also determine the stability. If r_{AB} is much larger than both r_{aB} and r_{Ab}, then cell type x_{AB} will out compete the other cell types regardless of any cooperation. This is reasonable since for any linear cooperation function with a fixed α, there will be an r_{AB} which is large enough that the cooperation function does not improve the growth of the other cell types enough to compete. If r_{AB} is not much larger than either r_{Ab} or r_{aB}, there can be a coexistence. This is reasonable since the cooperation functions depend on the cell type x_{AB}, and so cell type x_{Ab} or x_{aB} can be helped by the presence of x_{AB} to the extent that it is able to coexist with cell type x_{AB}. The lower left quadrant is split between two steady states. When the difference in growth rates between the once-mutated and twice-mutated populations can be made up by the cooperation functions, then there is a possibility of coexistence of all three cell types. However, it is also possible that the cooperation functions can make up for more than the

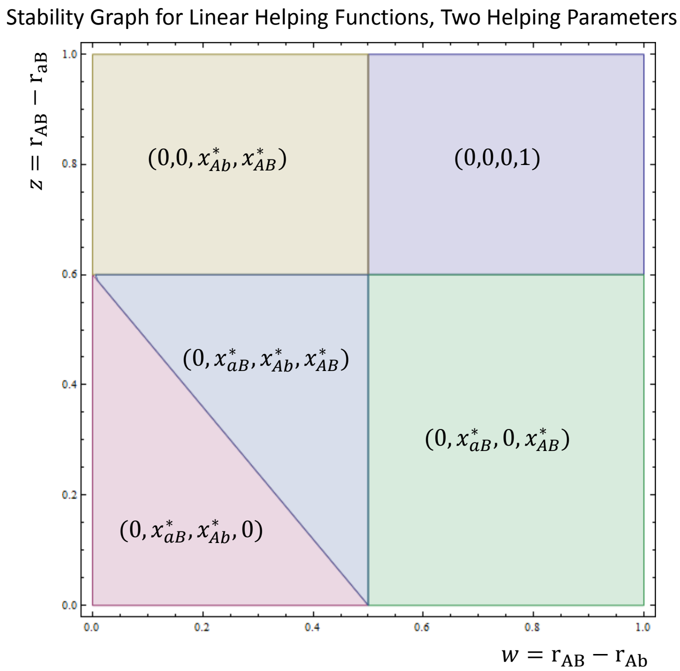


Figure 2.4: Graphs of the stability regions without mutations for symmetric cooperation with $\alpha = .6$, $\beta = .5$, assuming $r_{AB} > r_{ab}$.

difference even without the presence of x_{AB}, which would lead to coexistence of the two cell types x_{aB} and x_{Ab}.

2.2.3 More General Linear Cooperation Functions, Four Populations

The linear case examined before has one parameter to describe how cell type aB is assisted by the other cells in the environment and one for how cell type Ab is assisted, neglecting that the assistance by different cell types may in fact be different. We will consider a more general linear model with $g_{aB} = \alpha x_{Ab} + \beta x_{AB}$ and $g_{Ab} = \delta x_{aB} + \gamma x_{AB}$ to allow for different assistance from the different types of cooperation. If we continue to assume that $r_{AB} > r_{ab} > r_{aB}$ and $r_{AB} > r_{ab} > r_{Ab}$, then the steady states are similar to the previous model, but the requirements for stability are slightly different. Again, for ease of representation of the

$(x_{ab}, x_{aB}, x_{Ab}, x_{AB})$	Stability Conditions
$(1,0,0,0)$	Unstable under our assumptions
$(0,1,0,0)$	Unstable under our assumptions
$(0,0,1,0)$	Unstable under our assumptions
$(0,0,0,1)$	$\beta < z$ and $\gamma < w$
$(0, \frac{\beta+w-z}{\alpha+\beta}, \frac{\alpha+z-w}{\alpha+\beta}, 0)$	$\alpha z + \beta w < \alpha\beta,\ \beta > \frac{\alpha(r_{aB}-r_{ab})}{r_{ab}-r_{Ab}-\alpha}$
$(0, \frac{\beta-z}{\beta}, 0, \frac{z}{\beta})$	$\frac{-(\delta-\gamma)z+\beta(\delta-w)}{\beta} < 0$, positive populations
$(0, 0, \frac{\gamma-w}{\gamma}, \frac{w}{\gamma})$	$\gamma(\alpha-z) < (\alpha-\beta)w$, positive populations
$(0, \frac{(\alpha-\beta)w-\gamma(\alpha-z)}{(\alpha-\beta)\delta-\alpha\gamma}, \frac{(\delta-\gamma)z-\beta(\delta-w)}{(\alpha-\beta)\delta-\alpha\gamma}, \frac{-(\delta+\alpha)z+\alpha\delta}{(\alpha-\beta)\delta-\alpha\gamma})$	$\gamma\alpha + \beta\delta > \alpha\delta$, positive populations
$(\frac{\alpha\delta-\alpha r_{ab}-\delta r_{ab}+\delta r_{aB}+\alpha r_{Ab}}{\alpha\delta}, \frac{r_{ab}-r_{Ab}}{\delta}, \frac{r_{ab}-r_{aB}}{\alpha}, 0)$	Unstable if populations are positive

Table 2.4: Steady states and their stability conditions for the system 2.5

steady states, we will use $w = r_{AB} - r_{Ab}$, $z = r_{AB} - r_{aB}$, and $r_{aB} - r_{Ab} = w - z$.

Our choice of notation (w and z) conveniently allows for the graphical representation of the stable steady states as shown in figure 2.5. We can see from the figure that for the first parameter set (top two images), we have a region of bistability, where two different steady states are possibly stable, based on initial conditions. For the second parameter set (bottom two images of figure 2.5), there is a different steady state, where all three mutated populations can coexist, which did not have a region of stability for the first parameter set.

2.2.3.1 General Cooperation Function

Until now, the cooperation function was modeled as a linear function. We would like to see which conclusions we can make when the form of the cooperation function is not as limited. To address this, general functions $g_{aB}(x_{Ab}, x_{AB})$ and $g_{Ab}(x_{aB}, x_{AB})$ are considered. We considered functions $g_{aB}(x_{Ab}, x_{AB})$ and $g_{Ab}(x_{aB}, x_{AB})$ which are 0 at $(0,0)$. We included this restriction because if at $(0,0)$ the function is not zero, we can incorporate the value at $(0,0)$ into the constant growth rates, r_{aB} or r_{Ab}, and define a new cooperation function which is zero at $(0,0)$. In addition, since we are assuming that the cooperation is beneficial, we will consider only functions which are increasing with respect to both arguments. Furthermore

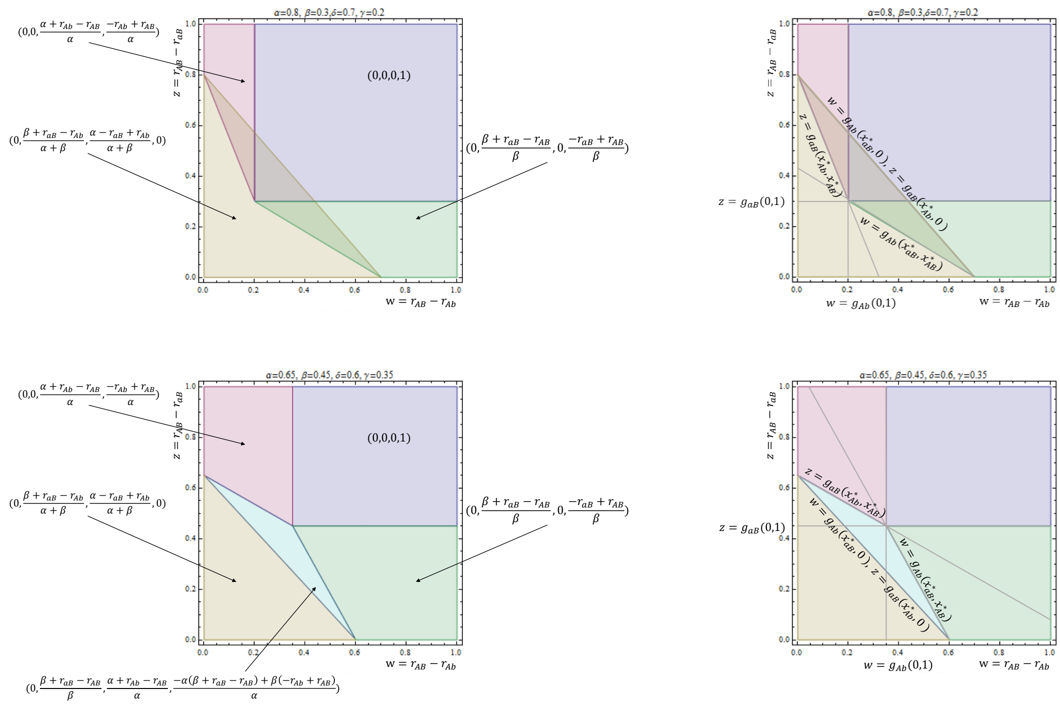


Figure 2.5: Graphs of the stability regions for cooperation functions of the form $g_{aB} = \alpha x_{Ab} + \beta x_{AB}$ and $g_{Ab} = \delta x_{aB} + \gamma x_{AB}$.

we will continue to restrict $r_{ab} > r_{aB}$, $r_{ab} > r_{Ab}$, $r_{AB} > r_{aB}$ and $r_{AB} > r_{Ab}$, but we will place no restrictions on the relationship between r_{AB} and r_{ab} or the relationship between r_{aB} and r_{Ab}.

Since it is not possible to solve for the fixed points without knowing the general functions g_{aB} and g_{Ab}, the values of the populations will be left as variables, rather than being expressed in terms of the parameters. We assume that the population profiles would follow the same form as the populations in subsection 2.2.2.2. We are able to make use of several constraints to simplify our calculations: cell populations must be positive, the differential equations must be equal to zero (since we are interested in the behavior of steady states), and the growth rates ($r_{ab},\ r_{aB},\ r_{Ab},\ r_{AB}$), cooperation functions ($g_{aB},\ g_{Ab}$), and first partial derivatives of the cooperation functions ($\frac{\partial g_{aB}}{\partial x_{Ab}},\ \frac{\partial g_{aB}}{\partial x_{AB}},\ \frac{\partial g_{Ab}}{\partial x_{aB}},\ \frac{\partial g_{Ab}}{\partial x_{AB}}$), are all positive for all positive values of the populations. We were subsequently able to determine constraints for the stability in the case with a general cooperation function. The constraints for the general case can be shown to be

$(x_{ab}, x_{aB}, x_{Ab}, x_{AB})$	Stability Conditions
$(1,0,0,0)$	$r_{AB} < r_{ab},\ r_{aB} < r_{ab},\ r_{Ab} < r_{ab}$
$(0,1,0,0)$	$r_{ab} < r_{aB},\ r_{AB} < r_{aB},\ R_{Ab}(1,0) < r_{aB}$
$(0,0,1,0)$	$r_{ab} < r_{Ab},\ r_{AB} < r_{Ab},\ R_{aB}(1,0) < r_{Ab}$
$(0,0,0,1)$	$r_{ab} < r_{AB}, R_{aB}(0,1) < r_{AB}, R_{Ab}(0,1) < r_{AB}$
$(0, x^*_{aB}, x^*_{Ab}, 0)$	$r_{ab} < R_{Ab}(x_{aB}, 0),\ r_{AB} < R_{aB}(x_{Ab}, 0),\ R_{aB}(x_{Ab}, 0) = R_{Ab}(x_{aB}, 0)$
$(0, x^*_{aB}, 0, x^*_{AB})$	$r_{ab} < r_{AB},\ R_{Ab}(x_{aB}, x_{AB}) < r_{AB}$
$(0, 0, x^*_{Ab}, x^*_{AB})$	$r_{ab} < r_{AB},\ R_{aB}(x_{Ab}, x_{AB}) < r_{AB}$
$(0, x^*_{aB}, x^*_{Ab}, x^*_{AB})$	$r_{ab} < r_{AB}$, equation 2.6
$(x^*_{ab}, x^*_{aB}, x^*_{Ab}, 0)$	Unstable if populations are positive

Table 2.5: Steady states and their stability conditions for the system 2.5. For notation purposes, we used $R_{Ab}(x_{aB}, x_{AB}) = r_{Ab} + g_{Ab}(x_{aB}, x_{AB})$ and $R_{aB}(x_{Ab}, x_{AB}) = r_{aB} + g_{aB}(x_{Ab}, x_{AB})$.

equivalent to the linear cooperation function constraints, in the case where the cooperation function is known to be linear. The steady states and their stability conditions are listed in table 2.5. For notation purposes, we will now use $R_{Ab}(x_{aB}, x_{AB}) = r_{Ab} + g_{Ab}(x_{aB}, x_{AB})$ and $R_{aB}(x_{Ab}, x_{AB}) = r_{aB} + g_{aB}(x_{Ab}, x_{AB})$.

$$\begin{aligned} \frac{\partial g_{Ab}}{\partial x_{AB}}(x_{aB}, x_{AB}) \frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab}, x_{AB}) + \frac{\partial g_{aB}}{\partial x_{AB}}(x_{Ab}, x_{AB}) \frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB}, x_{AB}) & \\ > \frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab}, x_{AB}) \frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB}, x_{AB}) & \end{aligned} \tag{2.6}$$

2.3 Conclusion

By starting with a simpler model, we were able to describe the behavior of system (2.1). In the most general case, we see that our model indicates stability of the steady states in which there is coexistence of all possible combinations of the three types of mutated cell types (x_{Ab}, x_{aB} and x_{AB}), depending on the parameters and initial conditions. However, the steady state of the unmutated cell type with the once-mutated cell types is not stable. In the context of cancer, this means that the normal cell types can not exist with the cancerous cells as a stable steady state.

This model supports coexistence of cooperators. We have stable steady states of just the specialized individuals (x_{Ab} and x_{aB}), as well as the specialized individuals and those that perform both tasks.

2.4 Application

The model discussed above is a more general version of the model introduced in [51], in which a system of quasispecies equations is used to model population dynamics when tasks are divided among the different cell types. There are 3 cell types which perform tasks: general, type 1 and type 2. The general type cell will perform all the needed tasks, while type 1 and 2 each perform a disjointed collection of tasks. In the model we will examine, it is assumed that a collection of individuals of size n will associate together to perform the two tasks, and only one individual capable of performing the task needs to be in the group in order for the task to be completed for the group. Therefore, the probability of successfully completing the task is based on the number of individuals in that group that can perform the task. So, in the case of the cooperation function that improves the growth rate of x_1, the probability that an individual can perform task 2 in a group of n individuals is $(1 - x_1^{n-1})$ (the first task is assumed to be performed as the cooperation is improving the growth rate of x_1). The maximum benefit gained from a task is b, and c_i are the costs associated with performing a task, both are assumed to be fixed constants. Following the quasispecies format of the previous systems, this system is then described by the system of equations:

$$x_G = r_G x_G - \phi x_G \tag{2.7a}$$

$$x_1 = r_1 x_1 - \phi x_1 \tag{2.7b}$$

$$x_2 = r_2 x_2 - \phi x_2. \tag{2.7c}$$

(x_1, x_2, x_G)	Stability Conditions
$(1, 0, 0)$	Unstable under our assumptions
$(0, 1, 0)$	Unstable under our assumptions
$(0, 0, 1)$	Unstable under our assumptions
$(x_1^*, x_2^*, 0)$	$r_1 + g_1(x_2, 0) > r_G$, $r_2 + g_2(x_1, 0) > r_G$, positive populations.
$(x_1^*, 0, x_G^*)$	Unstable under our assumptions
(x_1*, x_2^*, x_G^*)	Equation 2.9 and positive populations

Table 2.6: Steady states and their stability conditions for the system 2.7.

In system (2.7), the growth rates (including the cooperation functions) are described by the following functions:

$$r_G = b - c_1 - c_2 \tag{2.8a}$$

$$r_1 = (1 - x_1^{n-1})b - c_1 = (1 - (1 - x_G - x_2)^{n-1})b - c_1 \tag{2.8b}$$

$$r_2 = (1 - x_2^{n-1})b - c_2 = (1 - (1 - x_G - x_1)^{n-1})b - c_2. \tag{2.8c}$$

We can consider the cost as the growth rate, and $b - (1 - x_G - x_i)^{n-1}b$ as the cooperation functions:

$$g_1(x_2, x_G) = b - (1 - x_G - x_2)^{n-1}b$$

$$g_2(x_1, x_G) = b - (1 - x_G - x_1)^{n-1}b$$

In this way, we see that the system follows the form of the quasispecies model we examined. Based on the examination from section 2.2.2, we can easily see which states are stable and when. The steady states and the stability conditions of those steady states are detailed in table 2.6.

For the last steady state, we need a rather long condition:

$$\begin{aligned} &\frac{\partial g_1}{\partial x_G}(x_2, x_G)\frac{\partial g_2}{\partial x_1}(x_1, x_G) + \frac{\partial g_2}{\partial x_G}(x_1, x_G)\frac{\partial g_1}{\partial x_2}(x_2, x_G) \\ &\qquad > \frac{\partial g_2}{\partial x_1}(x_1, x_G)\frac{\partial g_1}{\partial x_2}(x_2, x_G) \end{aligned} \tag{2.9}$$

From the analysis of the stability conditions, we can have coexistence of the two specialized populations if group size is large or the costs associated with performing each task is large by the cooperation functions listed. Otherwise, a population of individuals capable of performing both tasks is stable. This aligns with the results found within [51].

Chapter 3

Cooperation Based Branching as a Mechanism of Evolutionary Speciation

3.1 Abstract

When performing complex tasks, coexistence of organisms in a shared environment can be achieved by means of different strategies. For example, individuals can evolve to complete all parts of the complex task, choosing self-sufficiency over cooperation. On the other hand, they may choose to split parts of the task and share the products for mutual benefit, such that distinct groups of the organisms specialize on a subset of elementary tasks. In contrast to the existing theory of specialization and task sharing for cells in multicellular organisms (or colonies of social insects), here we describe a mechanism of evolutionary branching which is based on cooperation and division of labor, and where selection happens at the individual level. The model we examine here also assumes that by performing multiple tasks, additional cost is incurred by the individual. Using a class of mathematical models and the methodology of adaptive dynamics, we investigate the conditions for such branching into distinct

cooperating subgroups to occur. We furthermore show that for a wide parameter range, this scenario is stable against the invasion of cheaters. We hypothesize that over time, this can lead to evolutionary speciation. Examples from bacterial evolution and the connection with the Black Queen Hypothesis are discussed. It is our hope that the theory of diversification rooted in cooperation may inspire further ecological research to identify more evolutionary examples consistent with this speciation mechanism.

3.2 Introduction

During their lifespans, many organisms engage in complex tasks associated with their various functions, such as nutrition uptake, interaction with the environment, and reproduction. There are different ways in which organisms undertake these tasks as a population. In one scenario, they may evolve in a way such that each organism can master all components of a complex task and perform them independently. In another, they can diversify such that distinct types of individuals undertake different complementary parts, and in some way share the products, each organism getting the full benefit but only performing a subset of tasks. The latter scenario can be facilitated by cooperation and division of labor. We would like to understand under which conditions such an evolutionary scenario may happen.

There are many instances of organisms that engage in division of labor rather than performing all the components of a complex task. Bacteria work together is biofilms, where cells attach to each other and to a surface, and secrete certain polymeric substances, which are a shared resource that can be used by all [52, 5]. Division of labor was also reported in bacteria *Pseudomonas fluorescens* [26] in which division of labor was observed to evolve, and mutants of a certain type emerged and cooperated with the parent strain to gain new territory. The two strains self-organized in space, with one providing a wetting polymer at the colony edge, and the other lagging behind and propelling the colony forward.

In cyanobacteria, strains *Prochlorococcus* and *Synechococcus* are thought to engage in cooperative behavior. They share a common ancestor which has the defense gene which encodes catalase-peroxide, (*katG*), capable of defending against external hydrogen peroxide. This gene remains present in *Synechococcus* and is missing in *Prochlorococcus*, yet both strains are sensitive to hydrogen peroxide. It is suggested that *Prochlorococcus* is able to take advantage of the other members of the community which do remove the hydrogen peroxide from the environment [37].

Another example of cooperation and division of labor in bacteria comes from the bacteria *B. subtilis*. The biofilm matrix of *B. subtilis* is primarily composed of two components: an exopolysaccharide (eps) and the protein TasA. To illustrate the fact that resources are shared between the bacteria, mutants of *B. subtilis* incapable of producing the eps and mutants incapable of producing TasA were studied. These genetically different mutants are able to colonize root systems together, but are unable to do so alone, indicating a sharing of their resources [2]. Thus, the cells have the ability to share resources extracellularly. It is thought that the cells that do not participate in the production of the matrix do not simply benefit from the work of the other cells, but are also involved in other processes that benefit the community [33].

In [29], a two-membered culture consisting of *Burkholderia cepacia* and *Stenotrophomonas maltophilia* is studied. When the organisms are growing on dodecyltrimethylamine as the sole source of carbon and energy, the two species engage in a commensalistic relationship; if however nitrogen-limited conditions are employed, the two enter a mutualistic relationship. In this case, *B. cepacia* only grows in the presence of *S. maltophilia*, which provides ammonium, and growth of *S. maltophilia* depends on the release of dimethylamine by *B. cepacia*. This is only one of the examples of many mutualistic relationships described in the literature studying mixed cultures with specific vitamin requirements [24].

There are further examples that suggest a connection between cooperation and diversification. In [41], evolutionary diversification of *Pseudomonas fluorescent* in soda glass vials is studied. In the course of about 10 days (less than 100 generations) a complex, interacting system of different bacterial types evolves, where genetically distinct types interact among each other in a variety of ways. In particular, three types are reliably identified: smooth morph (SM), wrinkly spreader (WS), and fuzzy spreader (FS). Most relevant for our study, the interactions between SM and WS types are of facilitation type; cooperating groups are formed by over-production of an adhesive polymer [42]. Further studies of diversifying selection were performed, which examined the positive effect that diversification could have in a cooperating system in the context of bacterial biofilms [6].

In [46], an artificial model of cooperation has been synthesized by using two genetically engineered strains of yeast, both obtained from *Saccharomyces cerevisiae*, by introducing a small set of genetic modifications. The two resulting strains have different metabolic capabilities and behave essentially as two different species, whose survival relies on cooperation with each strain supplying an essential metabolite to the other strain. This system was shown experimentally to be viable over a wide range of initial conditions, with an oscillating population ratio settling to a value predicted by nutrient supply and consumption.

A very different set of examples comes from the behavioral studies of higher organisms. It relies on the fact that behaviors can be taught and passed on to the next generation. In humans, trades are passed down through apprenticeships, and skill specialization is present in the different careers people choose. The division of labor has long been studied in economics; Adam Smith's book originally published in 1776 stated that "reducing every man's business to some one simple operation ... necessarily increases very much the dexterity of the work man" [47]. A recent example of specialization is from the medical field. As noted in [35], physicians are commonly expected to provide both ambulatory (outpatient) care and hospital care, but more recently, a group of physicians labeled as "hospitalists" focused their efforts on just hospitalized patients. In the work by Meltzer, it was found that as the costs of coordinating dropped, and as the cost associated with switching from hospital care to ambulatory care increased, the use of hospitalists rose. Together, the ambulatory care physicians and the hospitalists are able to provide the needed care for the population through specialization. Although this example seems very different from the example of cooperation on the bacterial world, there are certain similarities that are important for the present study. We are interested in cooperative strategies that are inherited (such as inherited genetic mutations in bacteria, or imitated behaviors in human social dynamics), and are under selective pressures that come at the level of individuals (and not at the level of colonies).

In the last decades, questions of specialization, cooperation, and division of labor have been studied extensively in the context of multicellular organisms, where a fundamental issue is the germline-soma divide, see e.g. [8], as well as cyanobacteria, which can serve as a natural model system for understanding differentiation and multicellularity [44]. In the latter paper, possible evolutionary paths leading to terminal differentiation in cyanobacteria are studied by examining different strategies under different environmental conditions. Paper [18] investigates how the division of labor was achieved in the face of selfishness of lower-level units. A mathematical model is introduced that describes the evolutionary emergence of the division of labor via developmental plasticity, focusing on viability and fertility of cells within a colony. In [45], a very general mathematical model is developed that identifies the exact conditions under which specialization arises, which can be applied to a great variety of situations that involve the existence of functionally specialized modules, including the evolution of specialized cell types, limb diversification in arthropods, and division of labor in social insects.

The theory developed in the present paper is fundamentally different from the works described above, because of the different levels of selection assumed. In papers [44, 18, 45], selection acts at the level of the cell colony, or the multicellular organism. Therefore, it is the fitness of the whole collection of cells that is being maximized. In the present paper we assume that different cells are in fact in competition among each other, each trying to maximize its own fitness. We further assume the existence of a certain level of mutualism or cooperation among the individuals in the population, and propose that division-of-labor type interactions comprise a mechanism of diversification, giving rise to evolutionary divergence and creation of different species.

In this paper we use the adaptive dynamics approach to study the dynamics of inherited traits and investigate conditions under which branching of the trait values occurs. In this context, branching means that a population monomorphic with respect to a given trait value becomes multiple monomorphic populations with different trait values, and division of labor among organisms performing different sub-tasks becomes the preferred evolutionary solution. The works that are most closely related to the present research and provided the theoretical basis for our developments are papers by Doebeli and colleagues [16]. Through the use of adaptive dynamics [12, 15, 16, 3, 34, 11, 13], we explore a model which displays diversification through cooperation.

3.3 Model

3.3.1 General formulation

Suppose in a population of individuals, reproductive fitness is conferred on those individuals who utilize results of a certain complex task. The completion of the task depends on fulfilling both of two sub-tasks, A and B, each of which requires a different set of skills. We assume that the tasks can be performed to various degrees. Consequently, an individual can be characterized by the ordered pair (a, b), where a expresses the investment of the individual towards the performance of task A, and b the investment towards the performance of task B. Since a and b are not percentages of investment, they are not bound above by one, but since they do represent an amount of investment to the population, a and b are bound from below by zero. Each task involves a certain cost (e.g. the time and energy involved in the fulfillment of the task). Moreover, individuals that perform both functions incur a cost associated with the diverse "toolkit" that they possess to be able to perform both A and B. Individuals that choose to specialize are able to avoid that cost. In our model, the extent of the ability to perform either of the two strategies is not directly related. An individual may perform both tasks, and the value of a is not dependent on the value of b, so the values may evolve independently. Based on these general assumptions, we can write down a general model for the payoff of a given individual:

$$\text{Payoff} = \begin{bmatrix} \text{Benefit if enough} \\ \text{of both products} \\ \text{are utilized} \end{bmatrix} - \begin{bmatrix} \text{Cost of} \\ \text{task } A \end{bmatrix} - \begin{bmatrix} \text{Cost of} \\ \text{task } B \end{bmatrix} - \begin{bmatrix} \text{Cost of} \\ \text{performing} \\ A \text{ and } B \end{bmatrix} \tag{3.1}$$

The different terms are explained below and specific examples are given.

Benefit. We assume that survival requires that the products from both task A and B must be available to an individual. We further assume that an interacting individual can share its products for task A and/or B, either through cooperation, or simply by means of a "leaky" function (that is, when the products automatically become available for others by means of, e.g., diffusion). Mathematically, an individual benefits by meeting two thresholds. An individual must have access to each product through either its own behavior or a combination of its own behavior and the the behavior of other individuals, which share resources. We further assume that the benefit is a growing and saturating function of the product amount. These requirements can be represented through the product of two sigmoid functions, which we denote as $S_{h,m}(x)$. The parameter h is the location of the threshold indicating the amount of product that is considered "sufficient" for survival, and m controls the sharpness of the threshold, with a larger m corresponding to a steeper curve. For example, we could have the

Hill function

$$S_{h,m}(x) = \frac{x^m}{x^m + h^m}.$$

Other possibilities include $S_{h,m}(x) = 1 + \tanh(m(x-h))$, $S_{h,m}(x) = \pi/2 + \arctan(m(x-h))$, and $S_{h,m} = \frac{1}{1+e^{-m(x-h)}}$. In this paper, unless indicated otherwise, the numerical examples use the Hill function. In Appendix B.1.1, we present an example using the arctangent.

Consider an individual, (a, b), in a monomorphic population (α, β). The function we used to describe the benefit the individual can expect to receive when interacting with any individual from the population is:

$$S_{h,m}(a+\alpha)S_{h,m}(b+\beta). \tag{3.2}$$

In this function, we assume full mixing of products and equal division of the product. An alternative formulation in the absence of saturation is examined in Appendix B.1.2; this model does not exhibit interesting behavior.

Two types of cost. Each individual pays both a cost associated with the performance of each task, and an additional cost of performing both tasks. The cost of performing an individual task should be a growing function of the investment, for example,

$$\delta(a+b)$$

for an individual (a, b), where δ is a constant. This linear cost could be considered the amount of time spent doing a specific task, or the cost of maintaining the functionality to perform the task. Unless otherwise indicated, this function was used in the simulations. In Appendix B.1.3 we also consider a quadratic function, $\delta(a^2+b^2)$, which exhibits qualitatively similar behavior.

Each individual also pays a cost for performing both tasks. The magnitude of this cost depends on the extent to which both tasks are performed. This cost is different in nature from the individual task cost described above, because it is only incurred if a sufficient amount of both products is being made. To give an example, making 1 unit of the first product and 99 units of the second product should be significantly easier than making 50 units of the first product and 50 units of the second product, because in the second case, an organism needs to keep up functional machinery to maintain both tasks. There could be a number of ways to express this idea mathematically. Unless otherwise indicated, in our simulations we used the product of two Hill functions,

$$\gamma S_{j,n}(a)S_{j,n}(b),$$

where γ is a constant. If both tasks are performed past the threshold j, then the cost will be relatively high. If at least one of the tasks is performed below the threshold, the cost is significantly smaller. Therefore, performing both tasks incurs a large cost, while performing only one does not. This type of cost structure is reasonable under the biological assumptions that there could be additional costs for maintaining the ability to perform two behaviors, or costs associated with the effort required to switch behaviors. Alternatively, one can use other formulations with such properties; in Appendix B.1.4 we showed that the simple product function γab also works.

Payoff and dynamics. The payoff an individual with behavior (a, b) in a population exhibiting behavior (α, β) is the sum of the benefits and costs listed above. To summarize the above description, we present the formulation of the payoff function that was used in most simulations:

$$P(a, b, \alpha, \beta) = S_{h,m}(a + \alpha)S_{h,m}(b + \beta) - \delta(a + b) - \gamma S_{j,n}(a)S_{j,n}(b). \tag{3.3}$$

Parameter	Description
δ	Coefficient of cost of individual task performance
γ	Coefficient of cost associated with performing both tasks
h	Threshold value for access to enough of each task
j	Threshold value for an individual's investment to a single task
m	Inverse width of Hill function associated with having access to enough of each task
n	Inverse width of Hill function associated with an individual's investment to a single task

Table 3.1: Model parameters and their descriptions.

Parameters of this model are summarized in table 3.1.

The payoff for an individual determines its likelihood to reproduce. In our simulation, the payoff of one individual with a number of partners is compared to another individual with a number of partners, and the individual with the higher payoff passes on its behavior to a new individual, with a small chance of mutation as described below.

The behavior is examined both numerically, by stochastic individual based simulations, and through the tools of adaptive dynamics. We will focus our model within the assumptions that we have a large, well mixed population. We are interested in whether this population can be invaded by a mutant individual with a trait different from the surrounding population. The fitness of the invader is compared to the population to determine if the invader is able to outperform the existing population, leading to a shift in the population behavior, or if the existing population is able resist the invader. As done in [13, 16], we will consider a payoff function, comprised of both a cost function and a benefit function, to determine the fitness of an individual in the surrounding environment.

3.3.2 Stochastic simulations

Individual-based stochastic simulations in our paper are inspired by the method described in [16]. We assume a fixed population of $1,000$ individuals to study the dynamics based on the given payoff function. At each iteration one individual, which we call "the focal individual", with trait values (a, b), is chosen to be removed and replaced either with its own offspring or with offspring of another individual, which is determined according to the individual's fitness (defined by the payoffs, as described below). To implement this, we compare the focal individual with another individual, a competitor (trait values (α, β)). The performances of the focal individual and the competitor are determined by the sum of payoffs obtained from pairwise interactions with all the other individuals. If the sum of the payoffs received by the focal individual when cooperating with the population is higher than the sum of the payoffs received by the competitor when cooperating with the population, then the focal individual's trait profile, (a, b), is used to replace the focal individual's traits, (a, b), with a small chance of mutation. Otherwise, the competitor's trait profile, (α, β), is used to replace the focal individual's traits, (a, b), with a small chance of mutation. The chance of a mutation occurring is 5%, and when a mutation occurs, each of the individual's trait values are perturbed by a normally distributed random number with mean zero and a standard deviation of .01. To prevent negative trait values, the larger between the new trait value and zero is selected to be the new trait value.

There are a couple of differences between our algorithm and the one used in [16]. For each fitness update, rather than updating proportional to the differences between payoffs, we updated based on the larger payoff function alone. Further, we used the sum of payoffs obtained from pairwise interactions with all the other individuals to determine the fitness value, whereas in [16], only a single individual is used. It is interesting that the number of interactions used to determine the fitness can have both quantitative and qualitative affects on the trait value evolution. These differences are illustrated in Appendix B.3. While all the simulations (unless otherwise stated) use the entire population to determine the fitness, we have determined that the central results of this study are not the consequences of the choice of the number of interactions. The branching behavior reported here can be observed under any number of interactions. The exact quantification of the effect of the number of interactions is subject of ongoing work.

Finally, we note that the methodology implemented here (and in [16]) is somewhat different from that in [12], where a non-constant population birth-death process is described. The relationship of the two approaches is much like the relationship between the Moran and Wright-Fisher processes on the one hand, and the contact process on the other. The non-constant population models (the contact process and the model of [12]) in statistical equilibrium possess similar properties to the constant population processes, see [27] for a discussion of this.

3.4 Adaptive Dynamics of Speciation

3.4.1 Theoretical considerations

The papers introducing the theoretical basis for adaptive dynamics were introduced in the 1990s [36, 12, 19] and adaptive dynamics was more recently summarized in The Hitchhiker's Guide to Adaptive Dynamics [4], which covers single trait evolution for Adaptive Dynamics. Our model uses multivariable traits, and the theory used in this paper comes from [31] and [32], which discuss the multivariable extension of adaptive dynamics. Through these methods, we can predict the behavior of the population.

We will be using a fitness function F which describes the fitness of an invader into a monomorphic population, similar to [16]. For our model, the fitness of an individual with trait vector

(α, β) in a population with trait vector (a, b) is the payoff of the individual less the payoff of the surrounding population:

$$F(\alpha, \beta, a, b) = P(\alpha, \beta, a, b) - P(a, b, a, b). \tag{3.4}$$

The gradient of this function is called the selection gradient. The selection gradient when two traits are being evolved (our case) is defined as the pair of functions

$$F_a(a, b) = \left.\frac{\partial F(\alpha, \beta, a, b)}{\partial \alpha}\right|_{\alpha=a,\beta=b} \tag{3.5a}$$

$$F_b(a, b) = \left.\frac{\partial F(\alpha, \beta, a, b)}{\partial \beta}\right|_{\alpha=a,\beta=b} \tag{3.5b}$$

where F is the fitness from equation (3.4). Then the change in each trait value follows the same formula. Assuming the absence of mutational covariance, we will examine this using trait a, although the formulation would be the same for b. The equation describing the dynamics of trait a is

$$\frac{d}{dt}a = k_a F_a(a, b) \tag{3.6}$$

where k_a is a non-negative coefficient that scales the rate of evolutionary change [12].[1] Here we assume that mutations are independent of the parameter values, so k_a is a constant. The steady states of the system with equations of the form of equation (3.6) are the singular strategies of the adaptive dynamics system. The stability of a steady state characterizes the initial convergent stability of the trait values. While equation (3.6) only predicts the behavior for monomorphic traits, we are interested in the evolution of a collection of individuals, all with their own traits. Realistically, the traits a of the population are not limited to a single value, and branching of the traits may occur. By branching, we mean that rather than the

[1] We should note that equation (3.6) may not necessarily always accurately predict the convergence of a stochastic simulation. This phenomenon is discussed in [30] in which so-called Darwinian demons can cause a simulation to not converge to the nearby stable singular strategy.

entire population maintaining the same trait vector, the population may split into multiple subpopulations each with different trait vectors. To determine how the trait values will behave, we first analyze the system by finding the convergent singular strategies by setting the selection gradients equal to zero. To determine the convergence stability, we examine the eigenvalues of the Jacobian of the selection gradients at the singular strategies. The Jacobian of the selection gradients evaluated at the singular strategy (a^*, b^*) is:

$$\left.\begin{bmatrix} \frac{\partial F_a(a,b)}{\partial a} & \frac{\partial F_\alpha(a,b)}{\partial b} \\ \frac{\partial F_b(a,b)}{\partial a} & \frac{\partial F_\beta(a,b)}{\partial b} \end{bmatrix}\right|_{a=a^*,b=b^*}. \tag{3.7}$$

If the real parts of the eigenvalues of the Jacobian (equation (3.7)) are not both negative, the singular strategy is either half stable or unstable. If both eigenvalues are negative, the convergent singular strategy is said to be convergent stable, meaning that the population is expected to initially converge to the singular strategy.

After converging, the subsequent behavior is dependent on the signs of the eigenvalues of the Hessian at the singular strategy. The Hessian of the fitness equation, equation (3.4), evaluated at the singular strategy (a^*, b^*) is:

$$H(a^*, b^*) = \left.\begin{bmatrix} \frac{\partial^2 F(a,b,a^*,b^*)}{\partial a^2} & \frac{\partial^2 F(a,b,a^*,b^*)}{\partial a \partial b} \\ \frac{\partial^2 F(a,b,a^*,b^*)}{\partial b \partial a} & \frac{\partial^2 F(a,b,a^*,b^*)}{\partial b^2} \end{bmatrix}\right|_{a=a^*,b=b^*}. \tag{3.8}$$

The eigenvalues of equation (3.8) indicate if the singular strategy is at a maximum, minimum or saddle. If the singular strategy is at a maximum (eigenvalues are negative) then the strategy is evolutionarily stable. If the singular strategy is at a minimum (eigenvalues are positive), then there is disruptive selection which may still lead to branching [32, 20]. If the singular strategy is at a saddle, then it is capable of being invaded on two sides, so the singular strategy is not evolutionarily stable, and branching is expected. Specifically, there will be branching along the direction of the eigenvector associated with the dominant eigenvector of the Hessian.

Solving for the singular strategies requires setting both equations in (3.5) to zero, finding the roots, and analyzing the corresponding Jacobian and the Hessian. To facilitate these calculations, we wrote a program in *Mathematica*. To find the singular strategies, we compared the list of points generated by *Mathematica's* ContourPlot for each of the equations in (3.5) when they were set equal to zero. These zero-sets are called evolutionary isoclines [4]. If two points (one from each list) were found to be each other's nearest neighbor, then they were appended to a list. That list of pairs was then run through *Mathematica's* FindRoot as the initial conditions to find much closer approximations to the intersections. These new points were added to another list with any duplicates removed. This last list was used to calculate the eigenvalues of the Jacobian of the two selection gradients and the eigenvalues of the Hessian of the fitness equation (3.4). The signs of these eigenvalues were then used to assign coloring to the images produced. Results of the computations are presented in the next section where we analyze parameter dependence of the evolutionary dynamics.

Note that an alternative approach to study the dynamics of cooperating populations of individuals could be used to analyze the bifurcation dynamics of ODEs that describe the coexistence of several discrete types of individuals, which are restricted in their behavior to strategies $(0, 0)$, $(1, 0)$ and $(0, 1)$. This approach however may give misleading results as shown in Appendix B.4.

3.4.2 Speciation

We will now explore the possibility of different evolutionary trajectories in an evolving population of individuals characterized by the assumptions as specified above. We will start by exploring the effect of changing parameter γ, which is the cost of performing both tasks, see table 3.1. We will use equilibrium analysis described above to characterize and visualize the

Color	Description of singular strategy stability
Red	convergent unstable
Green	convergent and evolutionarily stable
Cyan	convergent stable and evolutionarily unstable (directional branching)
Black	convergent stable and evolutionarily unstable (disruptive selection)

Table 3.2: Color coding used for the description of the equilibria in subsequent figures.

singular strategy and to predict the evolutionary dynamics. Colors are used to classify the singular strategies as summarized in table 3.2. When a strategy is both convergent stable and evolutionarily unstable, there are two possibilities. If the Hessian has both positive and negative eigenvalues at the singular strategy, then directional branching occurs in the direction of the unstable eigenvector. Disruptive selection occurs when the Hessian has only positive eigenvalues at the singular strategy; in this case it has been shown that the populations will likely form two sub-populations with trajectories predicted to converge to values that lie along the eigenvector associated with the dominant eigenvalue of the Hessian [20].

It is instructive to examine the behavior of the system as parameter γ, the coefficient of cost associated with performing both tasks, is varied. Figure 3.1 illustrates an example where directional branching is observed for a range of parameters. In this figure, we plot evolutionary isoclines, that is, the one-dimensional sets where the selection gradients are zero. Their intersections indicate the location of singular strategies, since at these points both selection gradients are zero. When γ is small, there are two singular strategies: a convergent, evolutionarily stable strategy and a convergent unstable strategy, see figure 3.1(top left). In this regime, the population is not predicted to branch into specialized populations because there is a small cost associated with performing both tasks. As γ grows, specialized traits will have a higher payoff compared to the payoff when performing both tasks, incurring the cost of γ. The singular strategy moves closer to the origin, to minimize the cost of contributing to both tasks, see figure 3.1(top center). For these intermediate values of γ, branching is observed, because individuals are able to avoid the cost of γ by specializing while still having

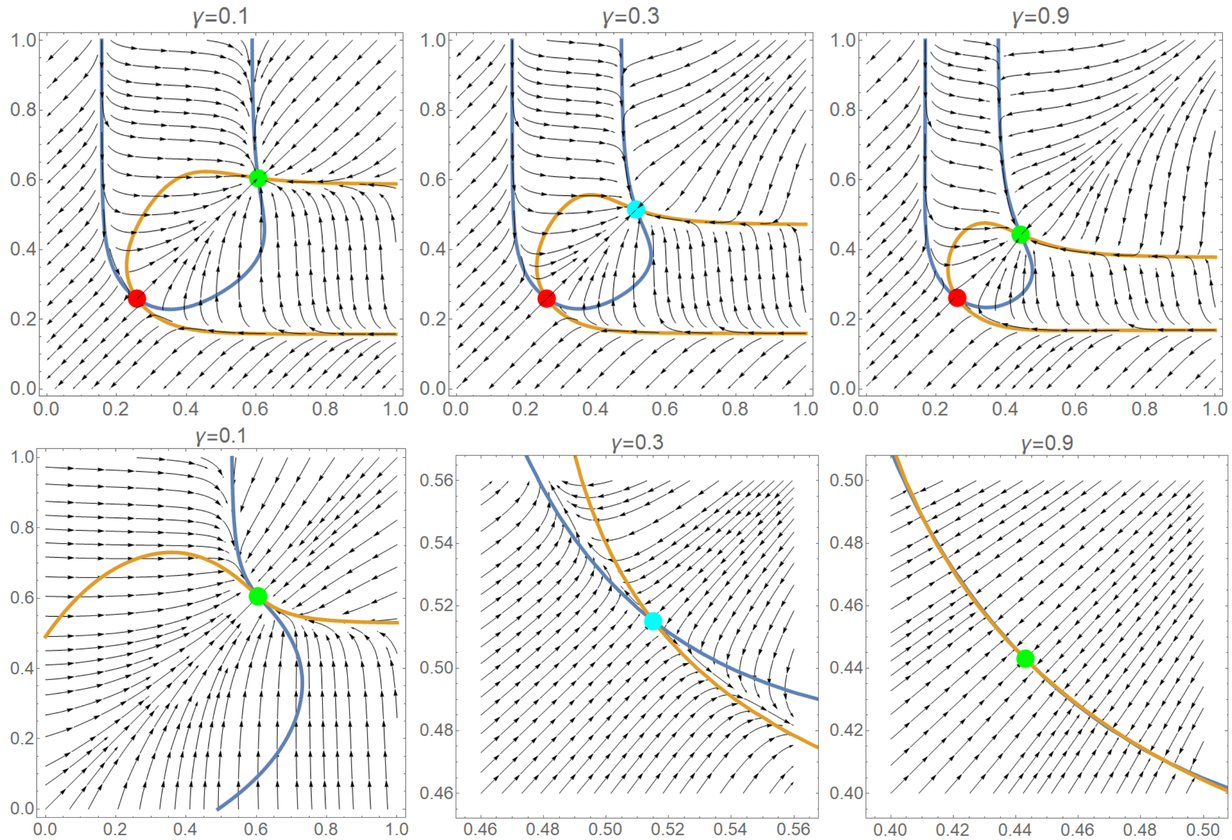
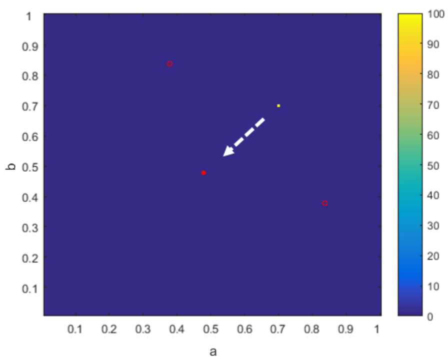


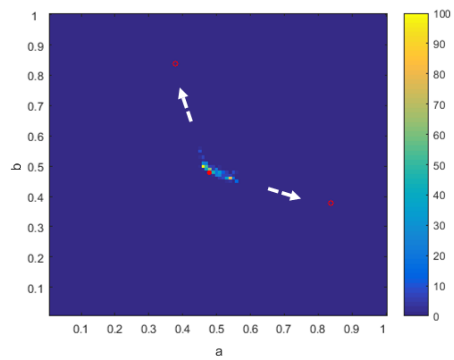
Figure 3.1: Top: Singular strategies, evolutionary isoclines and streamplot of the selection gradients as γ changes. The blue isocline is the derivative with respect to the first trait value, and the orange is with respect to the second. The stability of the singular strategies are color-coded per table 3.2. Bottom: Evolutionary isoclines and streamplot of the selection gradients for a mutant in an environment of individuals with the trait values of the convergent stable single strategies from the top, indicated with the color-coded dot. The rest of the parameters are $\delta = .1,\ h = .75,\ j = .5,\ m = 6,\ n = 6$.

access to enough of each product through cooperation with other individuals. Once γ grows large enough, any increase in the trait values negatively impacts the payoff function through the sigmoid functions, preventing specialization, see figure 3.1(top right).

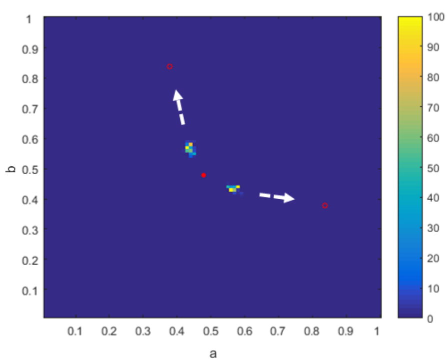
The top images of figure 3.1 show the behavior of a population of monomorphic individuals, while the bottom images display the streamlines and selection gradients of a single mutant in an environment of the convergent stable singular strategies indicated in the top row. The image at the bottom left shows that when the environment is comprised of individuals at the convergent stable singular strategy, the mutant will still gain payoff by also moving towards the convergent stable singular strategy. For the bottom middle, moving away from the environment's strategy results in improvement of the payoff. The bottom right indicates that the mutant would perform better along the evolutionary isocline shown; at a higher resolution (not shown here) one could see that the mutant would also perform better moving along that isocline towards the singular strategy.[2]



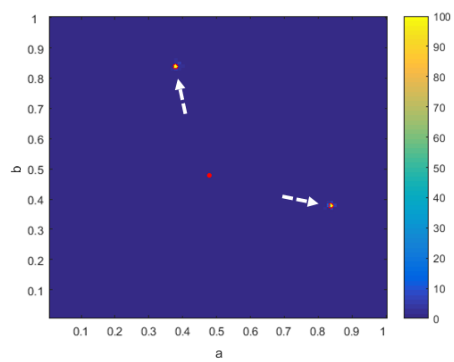
(a) Initial condition.



(b) Initial convergence.



(c) Branching.



(d) Convergence.

Figure 3.2: Stochastic simulations demonstrating evolutionary dynamics of the branching population. The density heat-plots show population distribution in the parameter space of the two tasks. The colorbar on the right of the images indicates the density of individuals. The four plots correspond to temporal snapshots. Convergence to a convergent-stable singular strategy is shown; then the population branches and converges to complementary singular strategies. The theoretical branching point and singular strategies of the two cooperating populations are indicated in red. The last figure has hollow circles to indicate the convergence for visibility of both the predicted points as well as the convergence of the simulation. The parameters are $\delta = .1,\ \gamma = .5,\ h = .75,\ j = .5,\ m = 6,\ n = 6$.

[2] For $\gamma = .9$, the dominant eigenvalue of the Hessian is -12.48 with eigenvector $(1, 1)$ while the non-dominant is only $-.089$ with eigenvalue $(-1, 1)$, indicating a much stronger influence towards the convergent singular strategy in the $(1, 1)$ direction, which is consistent with what is shown.

Numerical simulations were run to illustrate the branching behavior and to confirm that the theory aligns with the simulation observations, see figure 3.2. We started the simulations with an initial position centered at $(.85, .85)$ with a standard deviation of 0.01. Figure 3.2 presents a number of temporal snapshots, where the population of agents is represented as a density plot in the two-dimensional parameter space with the two axes being the individual's performance on the two tasks. Heat maps were created by partitioning the space into square bins and counting the number of individuals with parameter values in the bins. Thus, lighter colors correspond to higher densities and darker colors to lower densities. The first snapshot shows the initial population, centered around the symmetric point $(.85, .85)$. The red dot shows the analytically calculated convergent-stable point. In the second snapshot we can see the population converging to this point, and then splitting soon after (the third snapshot). The last two snapshots show the convergence of the two sub-populations to the convergent, evolutionarily stable points (for analytical predictions after the splitting see section 3.5).

As can be seen in the top images of figure 3.1 (and can be easily shown analytically), the origin in this case is an attractor. A population starting from $(0, 0)$ will not spontaneously evolve cooperation. This behavior around the origin however is irrelevant for the purposes of the present study. While extensive literature has been devoted to the origins and evolution of cooperation, in this paper we assume that cooperation is already in place, and explore the consequence of division of labor dynamics in the context of population branching. By modifying the local behavior of the payoff function near $(0, 0)$ one could change the stability properties of the origin; an example of this is worked out in Appendix B.1.3, where quadratic individual task costs can lead to the destabilization of $(0, 0)$. In this case, starting near $(0, 0)$, the population first evolves toward convergent stable/evolutionary unstable point characterized by division of labor, and then branches into two sub-populations. This however does not contribute meaningfully in our understanding of the evolution of cooperation itself. Therefore, in most models, we do not focus on the behavior at the origin but instead assume a population with division of labor as the initial condition.

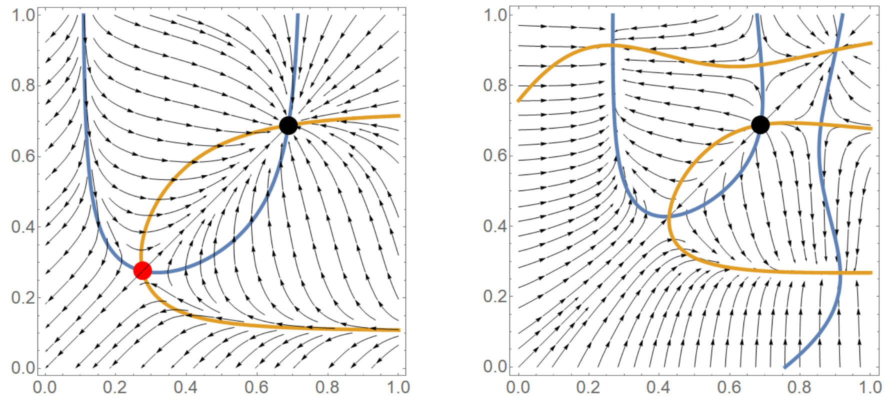
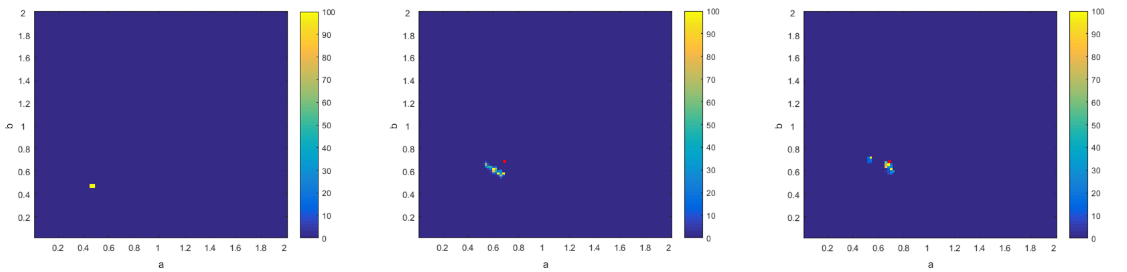


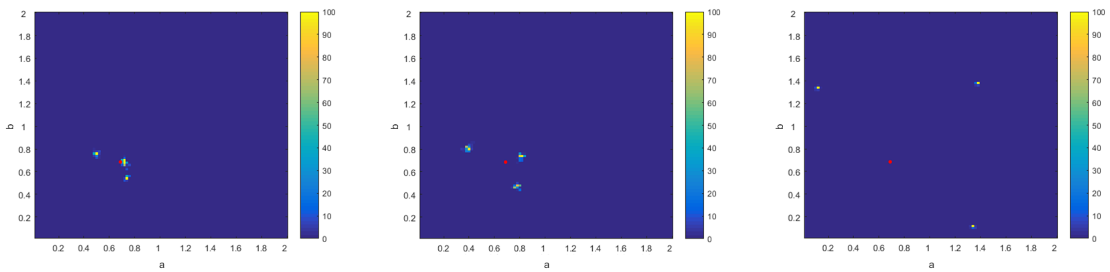
Figure 3.3: Left: Singular strategies, evolutionary isoclines and streamplot of the selection gradients for a monomorphic population. The stability of the singular strategies are color-coded per table 3.2. Right: Evolutionary isoclines and streamplot of the selection gradients for a mutant in an environment of individuals with the trait values of the convergent stable single strategy from the left, indicated with the color-coded dot. The parameters for both images are $\delta = .1, \gamma = .3,\ h = .8,\ j = .4,\ m = 3,\ n = 3$.

For the parameter sets examined so far in figure 3.1, the Hessian's eigenvalues at the singular strategies have either both been negative (convergent stable) or one has been negative and the other positive (which results in branching in the direction of the eigenvector associated with the positive eigenvalue). For the parameter set of figure 3.3, the Hessian's eigenvalues at the singular strategy are both positive. In this case, initially there should be there should be branching into a dimorphism in which the two clusters of behavior will lie along the eigenvector corresponding to the dominant eigenvalue [20]. For the parameters in figure 3.3, the eigenvector corresponding to the dominant eigenvalue is $(-1, 1)$, indicating population branching into specialists. To observe long-term dynamics, we used simulations. An example is shown in figure 3.4, where the population initialized with traits centered at $(.45, .45)$, exhibited disruptive selection, and subsequently split into three different trait profiles, which remained present and stable upon reaching the points indicated in figure 3.4 (f).

We observe specialization in this model with two traits. In Appendix B.2 we provide an example, which illustrates that branching in each trait independently is not a requirement to have branching behavior in a system with two traits.



(a) Initial Population Traits (b) Initial Convergence (c) Disruptive Selection



(d) Further splitting (e) Three Trait Profiles (f) Convergence

Figure 3.4: Stochastic simulations demonstrating evolutionary dynamics of the branching population, for the parameter set of figure 3.3. The density heat-plots show population distribution in the parameter space of the two tasks. The colorbar on the right of the images indicates the density of individuals. The six plots correspond to temporal snapshots after $0, 4 \times 10^4, 5 \times 10^4, 6 \times 10^4, 8 \times 10^4, 5 \times 10^5$ reproductions. Convergence to a convergent-stable singular strategy is shown; then the population branches and converges to three coexisting strategies.

3.4.3 Parameter variation

There are 6 parameters in the model, summarized in table 3.1. To examine the impact the different parameters have on the behavior, we varied the parameters other than δ and n to produce the images in figure 3.5. We fixed δ at .1, and n to be 3. We chose to fix δ because δ affects the location of the stationary point, but is not present in the Jacobian or the Hessian. There are two parameters, m and n, that vary the steepness of saturating functions. We chose to hold n constant for the images in figure 3.5, and vary m to capture how having different steepness between the two functions affects the behavior at the singular strategies. The parameter m was varied between 1 and 3, while the parameters j and h were varied from 0 to 1. To produce these images, we calculated eigenvalues of both the Jacobian and

Hessian, as described at the end of section 3.4.1 (this technique was also employed to create figures 3.1, 3.3). Each parameter set was color coded with the ratio between the components of the eigenvector of the dominant eigenvalue. From [20], if a dimorphic population profile can be sustained, the two branches should remain along that eigenvector. However, as shown in section 3.6 it may be that branches along other directions are the ones that persist long term.

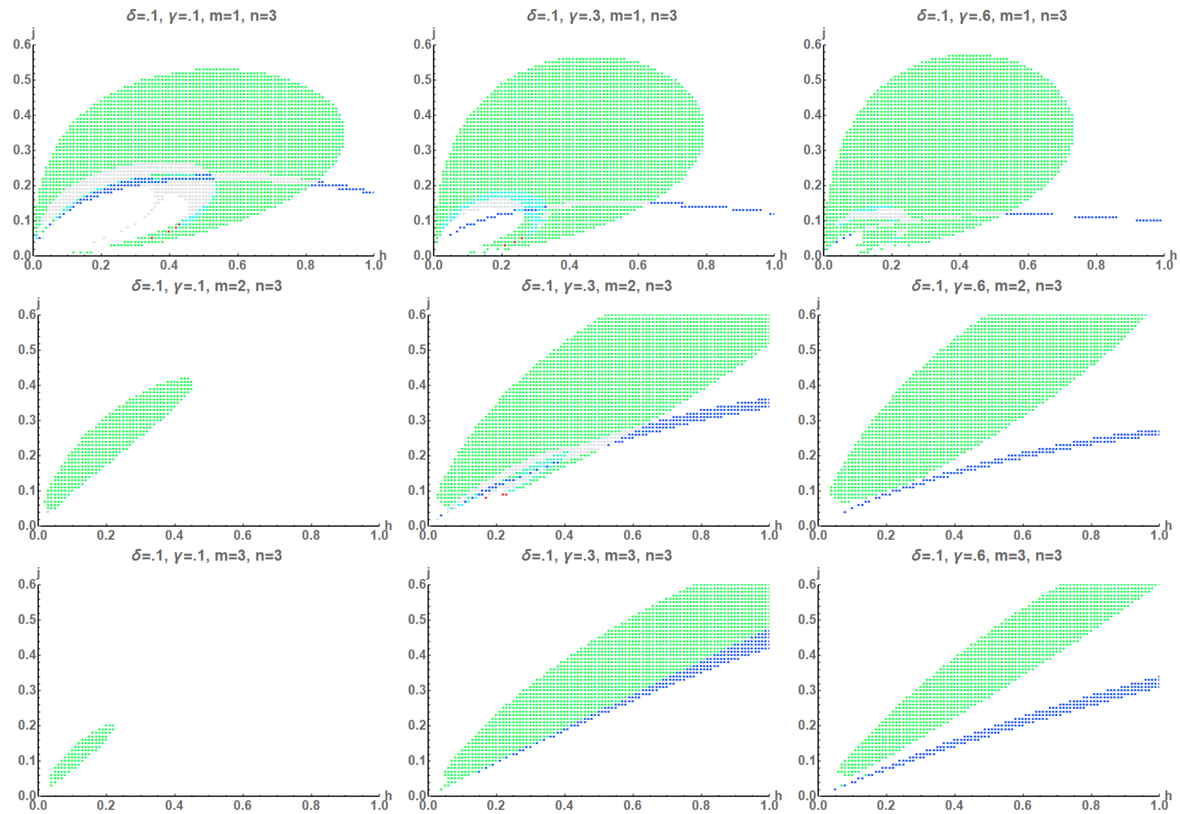
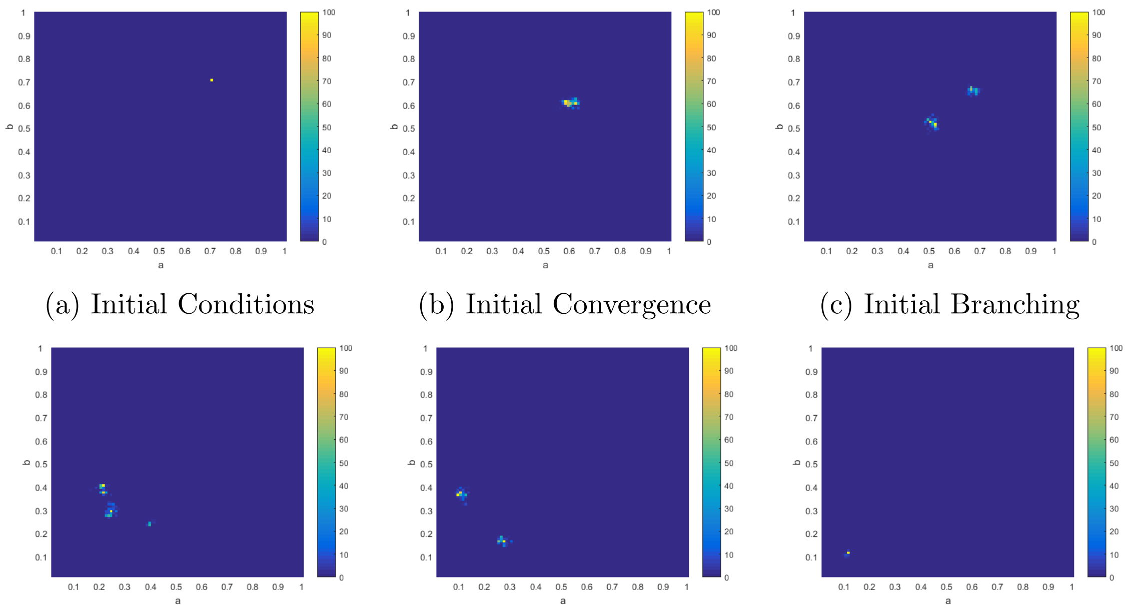


Figure 3.5: Parameter variation. From left to right, γ is varied through $.1, .3, .6$ and from top to bottom, m is varied through $1, 2, 3$. In each plot, along the horizontal axis is h from 0 to 1, and along the vertical axis is j, also from 0 to 1. Colors indicate the expected direction of branching. Bright green is along the $(-1, 1)$ direction with blue along the $(1, 1)$ direction. As the color changes from yellow to red, the eigenvector associated with the dominant eigenvalue is approaching $(-1, 0)$. The empty circles represent points that have more than one convergent singular strategy, that is also evolutionarily unstable and these different strategies may each experience branching in a different direction. We also note that branching along the $(-1, 1)$ diagonal is relatively unusual when compared to the larger set of parameters in which branching along the $(1, 1)$ direction occurs.

In figure 3.5, we see that for a wide variety of parameters we will see branching (any non-white point in the figure corresponds to a parameter set where branching is observed). For most of these parameter sets, we see branching along the $(-1, 1)$ vector, indicating the emergence of specialized populations that cooperate. In addition, from figure 3.5, we can see how the steepness values of the cost and benefit functions interact, as well as how the thresholds of the benefit and cost influence branching behavior. As the steepness of the benefit function grows, the region in which specialist populations will arise condenses to the region in which the thresholds are similar. If the threshold h is much smaller than j, then the population is able to reap the benefits of being independent while also avoiding the costs associated with production. If the threshold j is much smaller than h, and if the trait value of the population is close to the convergent singular strategy, then the majority of the cost γ is already being incurred, and lowering the trait value of a mutant only lowers the benefit received without lowering the cost very significantly. So, for most parameter sets, h and j must be relatively close to one another to see specialization. We can also see how γ affects the region in which we expect to see branching into specialists. As γ grows, the region in which specialization may occur grows. Analytically, this is because a greater cost can be avoided by specializing, which encourages branching into specialist groups.

It follows from figure 3.5 that there are parameter sets that show branching in the $(1, 1)$ direction, indicating the possibility that a population of defectors and a population of cooperators might emerge. Figures 3.4 and 3.12 however demonstrate that the direction of initial branching is not necessarily the ultimate direction that coexisting trait profiles settle to. In both figures 3.4 and 3.12 we see coexisting populations that both cooperate, and not the defecting and cooperating populations that would emerge if the population branched along the dominant eigenvalue's eigenvector, which is $(1, 1)$ in both of those cases.

There are parameter combinations that demonstrate branching along the $(1, 1)$ diagonal, as shown in figure 3.6. Under the parameter values in this example, we have one positive eigenvalue, with eigenvector $(1, 1)$, indicating branching along the $(1, 1)$ direction. Indeed, in figure 3.6, we see that there is initially branching along the $(1,1)$ vector, but it subsequently branches into two coexisting populations along the $(-1,1)$ vector, and eventually reaches the monomorphic singular strategy at $(.09699, .09699)$. So, although the branching did occur along the $(1,1)$ vector, this behavior was ultimately unstable and both populations could not be maintained.



(a) Initial Conditions (b) Initial Convergence (c) Initial Branching

(d) Second Branching (e) Two Populations (f) Convergence

Figure 3.6: Stochastic simulations demonstrating evolutionary dynamics of the branching population, for the parameter set $\delta = .1$, $\gamma = .1$, $h = .92$, $j = .2$, $m = 1$ and $n = 3$. The density heat-plots show population distribution in the parameter space of the two tasks. The colorbar on the right of the images indicates the density of individuals. The six plots correspond to temporal snapshots after 3×10^4, $5 \times 10^4, 8 \times 10^4, 11 \times 10^5, 13 \times 10^5, 2 \times 10^6$ reproductions. Convergence to a convergent-stable singular strategy is shown; then the population branches along the $(-1,1)$ diagonal, the population with smaller of trait values then branches again, the population with higher trait values dies out, and the remaining populations converges together to a single monomorphic strategy.

We have examined two types of situations in which branching along $(1,1)$ was predicted: when there are two positive eigenvalues and the dominant eigenvalue has eigenvector $(1,1)$ and when there is only one positive eigenvalue, which has eigenvector $(1,1)$. Neither of these situations resulted in the stability of a population of defectors and a population of cooperators. The intuitive reason for this (for the class of models described by equation (3.1)), is as follows. When branching happens initially along the $(1,1)$ diagonal, it creates two groups of individuals that do a similar thing: they are both "generalists", that is, they both participate in both tasks, but they do it to a different extent. Note that both groups are self-sufficient, and because of trait value differences, one performs better than the other. Therefore, depending on the parameters, we expect only one of these groups to survive. (This reasoning clearly does not apply in the case when branching in the $(-1,1)$ direction occurs. There, the two groups depend on each other for survival, leading to a qualitatively different scenario, which is the main focus of our study). So, only one of the generalist groups will remain, after which a number of scenarios are possible, including further branching along the $(-1,1)$ diagonal, as is the case in figure 3.6. Branching along the $(1,1)$ diagonal, however, is not the only way to have a rise in defectors. Another scenario is fully explored when we introduce a population of cheaters in section 3.6. There, one of the populations is initially at $(0,0)$; the two populations are qualitatively different, and they can coexist for a range of parameter combinations, as described below. This point however cannot be achieved by a gradual change starting from $(1,1)$ branching, but only by a "catastrophic" event, which changes the phenotype to (0,0).

3.5 Predicting behavior after branching

As explained in [19], it is possible to predict which values the population will converge to after the branching behavior occurs. We will apply the methods from adaptive dynamics, as used in section 3.4, and focus on the case where the population is predicted to branch along the $(-1, 1)$ vector.

Assume that a singular strategy of the form (α, α) is predicted to branch along the vector $(-1, 1)$. We can show that the population will split into two evenly sized branches. The populations will move away from the singular strategy (α, α) at the same rate so that the trait vectors will have the forms $(a+b, a-b)$ and $(a-b, a+b)$ due to the symmetry of the payoff function ($P(x, y, a+b, a-b) = P(y, x, a-b, a+b)$). Without loss of generality, let us call the fraction of the population with trait vector $(a+b, a-b)$ by N, and therefore the fraction of the population with with trait vector $(a-b, a+b)$ is $1-N$. The growth of N is described by

$$\begin{aligned}\frac{dN}{dt} = N\Bigg(&NP(a+b, a-b, a+b, a-b) + (1-N)P(a+b, a-b, a-b, a+b)\\ &-\Bigg(N^2P(a+b, a-b, a+b, a-b) + N(1-N)P(a+b, a-b, a-b, a+b)\\ &+ N(1-N)P(a-b, a+b, a+b, a-b) + (1-N)^2P(a-b, a+b, a-b, a+b)\Bigg)\Bigg).\end{aligned} \tag{3.9}$$

By setting $\frac{dN}{dt} = 0$, we find that there are three steady states of the ordinary differential equation: $N = 1$, $N = 0$, and $N = 1/2$. The steady state $N = 1/2$ is stable if the derivative with respect to N of the right hand side of equation (3.9) is negative. This is equivalent to

$$S_{h,m}(2a+2b)S_{h,m}(2a-2b) < S_{h,m}(2a)^2. \tag{3.10}$$

Note that when $b = 0$, we have $S_{h,m}(2a+2b)S_{h,m}(2a-2b) = S_{h,m}(2a)^2$, so we can determine the stability of $N = 1/2$ if we know whether $S_{h,m}(2a+2b)S_{h,m}(2a-2b)$ is increasing or decreasing as b increases from zero. If $S_{h,m}(2a+2b)S_{h,m}(2a-2b)$ is decreasing, then we have that $S_{h,m}(2a+2b)S_{h,m}(2a-2b) < S_{h,m}(2a)^2$, and $N = 1/2$ is stable. To this end, we will examine the derivative of $S_{h,m}(2a+2b)S_{h,m}(2a-2b)$ with respect to b (here we will assume the hill function for the form of the benefit function):

$$-\frac{4^m(a^2-b^2)^{m-1}h^m(2^m((a+b)^{1+m}-(a-b)^{1+m})+2bh^m)m}{((2a-2b)^m+h^m)^2((2a+2b)^m+h^m)^2}.$$

Assuming that $a+b$ and $a-b$ are both positive, as they are investment into a task, we can see that both the denominator and numerator are positive. Therefore, as b increases, $S_{h,m}(2a+2b)S_{h,m}(2a-2b)$ decreases, so $S_{h,m}(2a+2b)S_{h,m}(2a-2b) < S_{h,m}(2a)^2$, and therefore $N = 1/2$ is stable so long as the trait values are positive.

When branching occurs in the $(-1,1)$ direction, the average fitness of a population with two subpopulations expressing behaviors $(a+b, a-b)$ and $(a-b, a+b)$ is therefore

$$\frac{1}{2}(F(a+b,a-b,a+b,a-b)+F(a+b,a-b,a-b,a+b)). \tag{3.11}$$

Thus, we can find the convergent stable singular strategies of the new dimorphic population. As before, the singular strategies are found by setting the selection gradients of equation (3.11) equal to zero. We then determine the stability of those singular strategies by finding the eigenvalues of the Jacobian of the selection gradients. Lastly, the evolutionary stability is determined by finding the eigenvalues of the Hessian. The singular strategies that are both convergent and evolutionarily stable are the strategies that will be stable after the initial branching.

We illustrate these results by performing the analysis for parameters $\delta = .1$, $\gamma = .5$, $h = .75$, $j = .5$, $m = 6$ and $l = 6$, which is visualized both in figure 3.7, which shows the analytical predictions, and in figure 3.2, which shows the behavior of a stochastic simulation. Figure 3.7 indicates that the population is expected to converge first to the point $(.47929, .47929)$. Subsequently, evolutionary branching is predicted to occur, and the two cooperating populations will converge to task value profiles of (.377339, .838379) and (.838379, .377339). Numerical simulations of figure 3.2 show the convergence of the two sub-populations to the convergent, evolutionarily stable points, as analytically calculated (they are marked by two red dots). These dynamics are in close agreement with quantitative theoretical predictions.

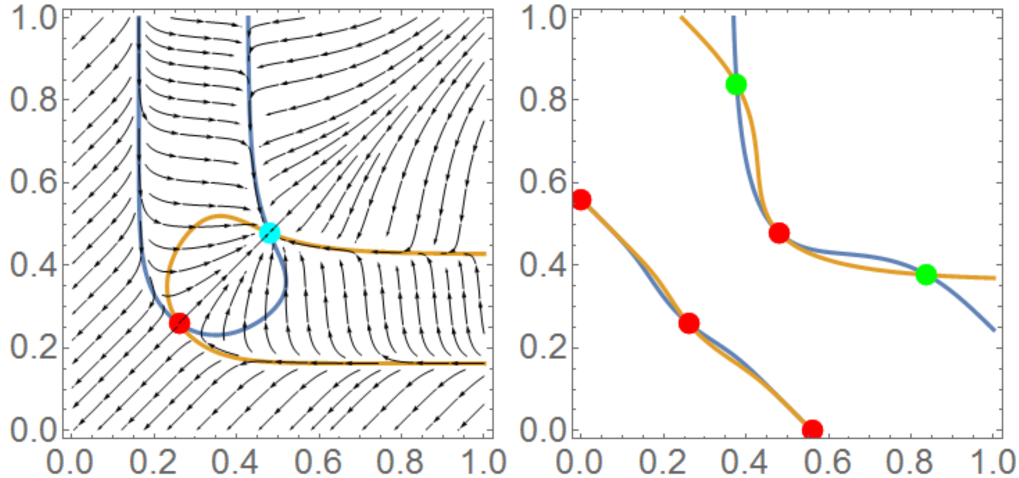


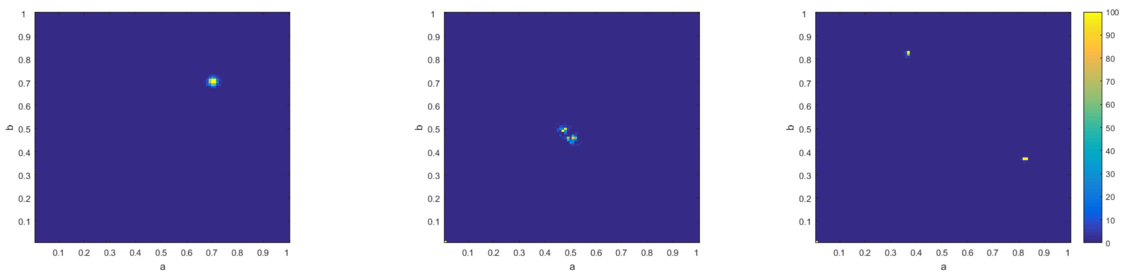
Figure 3.7: (Left) Singular strategies for $F(a, b, \alpha, \beta)$. (Right) Singular strategies for $F(a, b, \alpha, \beta) + F(a, b, \beta, \alpha)$. Phase plane indicators are omitted from the second image as there is no longer a single parameter, and the initial monomorphic population will first converge to the convergent stable singular strategy and branch along the unstable eigenvector in the direction $(1, -1)$. The parameters are $\delta = .1,\ \gamma = .5,\ h = .75,\ j = .5,\ m = 6,\ n = 6$.

3.6 Cheaters

In [39] the role of cheaters in a cooperative microbial community is studied and it is concluded that under a range of circumstances, cooperation may not be the best evolutionary strategy, because relying on a partner for fulfilling a task may be dangerous (because such a partner may not be available in the vicinity). This problem has been addressed theoretically in [28], which investigated under what conditions cooperating populations with cheaters can stably survive. In order to assess the influence of cheating in the present setting, that is, in the context of cooperation-dependent speciation, we investigate the dynamics of populations that contain individuals which do not contribute to the common goods, but still take advantage of the common goods provided by cooperators. In the simplest case, we consider an individual to be a cheater if the individual does not contribute to either task. Such individuals are represented in our model by the task profile $(0, 0)$. These individuals do not participate in resource production, but do benefit from the population's behavior since they are subject to the same payoff function. To examine how cheaters impact cooperation and diversification, we performed a number of simulations, as described below.

We started the simulations with a standard monomorphic population of individuals at symmetric task values: (a, a). Cheaters were introduced as a kind of a "mutation": the winning individual's offspring became $(0, 0)$ with a fixed probability ρ per reproduction. The simulation would then proceed as before: competition followed by reproduction and chances for mutations of the trait values, now with a chance of mutation into a cheater as well. A typical outcome is presented in figure 3.8. The system evolves as a relatively homogeneous population, and then splits into two groups with complimentary skills. The population dynamics is presented in figure 3.9. We can see a stratification of the population into cheaters and non-cheaters. The presence of cheaters can also be seen in figure 3.8 as a dot at point (0,0). It is observed that the population centered around $(0, 0)$ tends not to spread out. This is because although individuals of type $(0, 0)$ mutate at the same rate as the rest, the task parameters are forced to stay positive. Therefore, the population of cheaters will stay near the point $(0, 0)$ since they cannot decrease from $(0, 0)$, and increasing in any direction would punish an individual severely, as they would incur a cost and not gain much for their efforts. This is consistent with our previous results about the singular strategies: a population starting at $(0, 0)$ will not begin to participate in task performance due to the unstable singular strategy close to $(0, 0)$ which forces the population away from positive trait values.

The rate at which offspring could change to cheaters, ρ, was varied in further simulations. For the particular parameter set used in figure 3.8, if ρ was below 0.15 (which is a relatively high number), the cheaters did not disrupt the cooperative behavior seen previously. When the mutation rate was set to 0.2, the cheaters prevented branching, and when the mutation rate reached .29, the cheaters grew to take over the population. We conclude that the evolutionary scenario discussed in this study is reasonably robust against the invasion of cheaters, and persists under relatively high "cheating" mutation rates.



(a) Initial population. (b) Branching. (c) Convergence.

Figure 3.8: Stochastic simulation with cheaters converging to convergent-stable singular strategy, branching and then converging to complementary singular strategies. For this simulation, $\delta = .1$, $\gamma = .5$, $h = .75$, $j = .5$, $m = 6$, $n = 6$, $\rho = .15$ and $\mu = .05$.

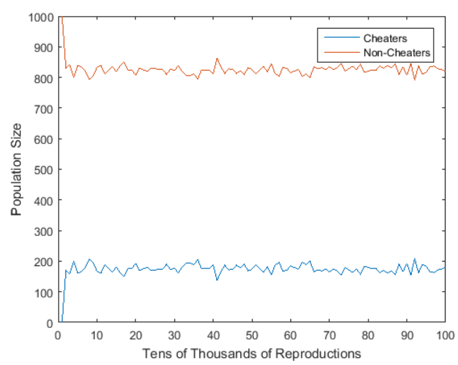


Figure 3.9: Population dynamics of cheaters and non-cheaters restricted to the first 10^6 reproduction events in the simulation corresponding to figure 3.8. The behavior remains similar for longer periods of time.

In a more detailed model of cheating, we explored the possibility of cheating in the first, the second, or both tasks. In this case, offspring of individual (a, b) could become $(a, 0)$, $(0, b)$, or $(0, 0)$ with probability $\rho/3$ each. In this scenario, branching was also observed, see figure 3.10, as long as the magnitude of probability ρ was below a threshold (for the parameters used in figure 3.10, we used $\rho = 0.1$). In figure 3.10 (left), the partial cheaters that do not perform/contribute to the first (second) task can be seen along the edges of the axes, corresponding to the branched populations. The population of full cheaters appears as a point around $(0, 0)$.

If the chance of "cheating" mutation is taken higher than a threshold, then the cheaters take over and the non-cheating population dies out. In principle, the remaining cheater population is not sustainable, and such dynamics will result in "evolutionary suicide" [21, 22]. This however was not pursued further in this study, because in our current models, the population is fixed at a constant size.

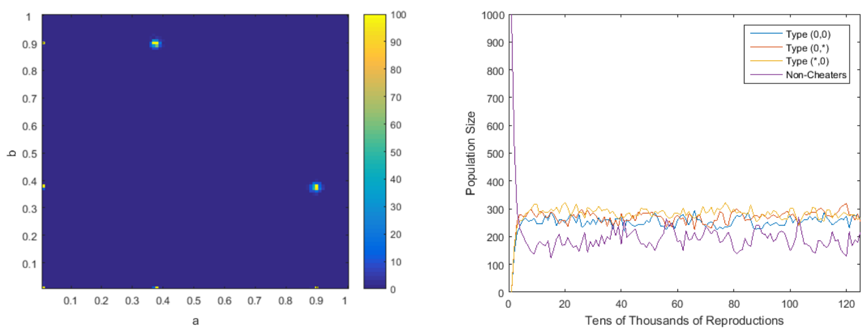


Figure 3.10: Inclusion of partial cheaters. (Left) Simulation after converging to complementary singular strategies. (Right) Population dynamics of cheaters and non-cheaters. Cheaters were categorized based on the task not performed. For this simulation, $\delta = .1$, $\gamma = .5$, $h = .75$, $j = .5$, $m = 6$, $n = 6$, $\rho = .1$ and $\mu = .05$.

So far we have observed cheaters to either not influence the branching dynamics qualitatively, or (if the "cheating" mutation rate, ρ, is higher than a threshold) displace non-cheaters and drive the population of non-cheaters extinct. There is a different scenario that can also occur. When cheaters are introduced, the population does not branch, but instead converges to a monomorphic singular strategy. The resulting population consists of both cheaters and non-cheaters, where non-cheaters perform both tasks extensively and the cheaters are sustained, see figures 3.13 and 3.14. In fact, the cheaters grow to a large portion of the population, but rather than driving the non-cheaters to extinction, a balance is formed between the two profile types. This result is similar to that obtained in [28], where we considered the crossing of a fitness value by means of consecutive mutations, in the presence of cooperation among the cooperative mutants and a cheater population. An advantageous multi-hit mutant was created faster in the presence of cooperation, and an intermediate, transient state was dominated by cheaters whose survival was dependence on the presence of cooperators, very similar to what is observed in the current model.

Finally we note that there are other ways in which non-cooperators can be included in the simulation. For example, one could envisage a mutation that enables individuals to produce both products but not share them. Such types can be assumed to pay the same price as the wild types for producing the product(s), but since they are not sharing them, then can utilize their own product to a larger extent, thus getting more payoff for producing less. In the present framework these types take over the population, either preventing branching from happening, or destroying the population after branching. This scenario, however, cannot happen in the case where division of labor is in place because of a "leaky" product, whose sharing the organisms simply cannot switch off.

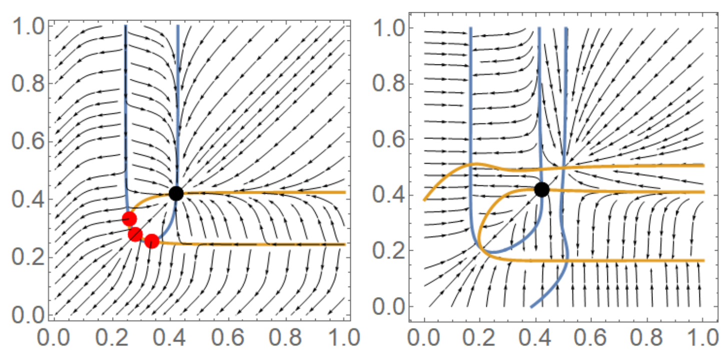
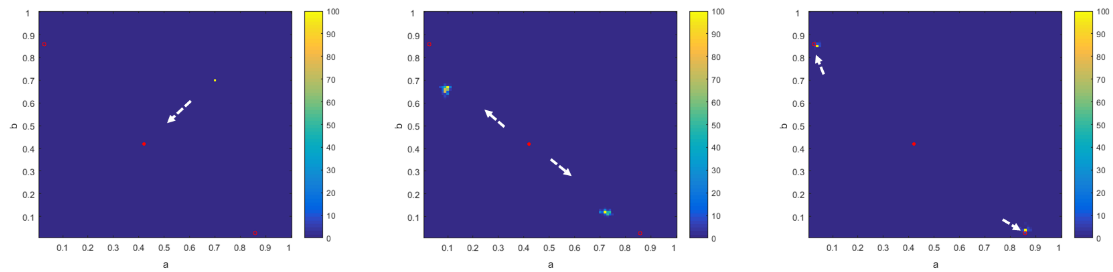


Figure 3.11: Evolutionary dynamics without cheaters for the parameter set $\delta = .1,\ \gamma = .5,\ h = .5,\ j = .23,\ m = 6,\ n = 6$. (Left) Isoclines, singular strategies and stream plot of the selection gradients for a monomorphic population, strategies are color-coded per table 3.2. (Right) Isoclines and stream plot of the selection gradients for a mutant trait value when the resident population is the convergent stable singular strategy from (left), shown here by a black dot.

3.7 Discussion

In this paper we propose a possibility that speciation can be facilitated by the presence of collaboration among individuals. If complex tasks are involved in the organisms functioning, and if products of specific sub-tasks can be shared, then in a wide class of models, evolutional



(a) Initial Population (b) Branching (c) Convergence

Figure 3.12: Simulation over 400000 updates for the system with the same parameters as in figure 3.11, without cheaters.

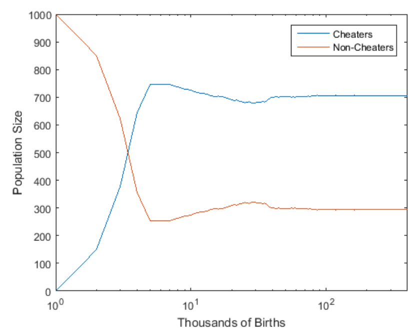
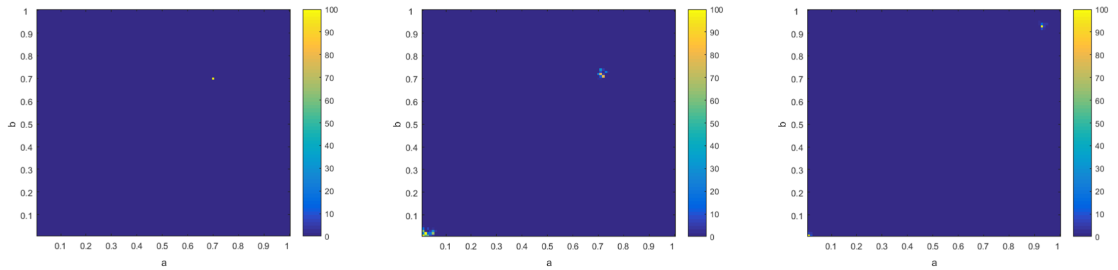


Figure 3.13: Population dynamics over 400000 updates to the trait values, in the presence of cheaters ($\rho = .1$) for the parameter set $\delta = .1,\ \gamma = .5,\ h = .5,\ j = .23,\ m = 6,\ n = 6$. See also figure 3.14.

dynamics of speciation is possible. The assumptions leading to the possibility of branching are that (i) the organisms receive an increased benefit from products of both (and not just one) sub-task, (ii) products are shared, and (iii) participating in both sub-tasks leads to additional costs compared to just one sub-task. Analysis shows under which circumstances a population may diversify, splitting into two subpopulations that each specialize on a certain sub-tasks. If speciation does not happen, the population remains homogeneous, and all individuals must complete all components of the task.

There are several biological examples that can be considered as indirect evidence of the mechanism of speciation described here, some of them are mentioned in the Introduction. In [26], a single strain of *Pseudomonas fluorescens* evolves into two genetically different strains, which participate in division of labor to better spread the colony outwards; the authors



(a) Initial Population (b) Growth of Cheaters (c) Convergence

Figure 3.14: Simulation after 400000 updates for the system with the same parameters as in figure 3.13, including a possible mutation into cheaters with probability $\rho = .1$.

suggest that one type pushes along the other, while the other secretes a lubricant to assist in the motility of the population. The mechanism of speciation is consistent with our model. The two types of cyanobacteria, *Prochlorococcus* and *Synechococcus*, engage in mutualistic interactions, and are thought to share a common ancestor [37]. Our model suggests a way in which *Prochlorococcus* and *Synechococcus* could have evolved into the two strains they are now without requiring a barrier between the populations. In [2], two strains of *B. subtilis* were studied, which cooperated to enable the population to colonize root systems (while each strain was incapable of doing so alone). In [41], *Pseudomonas fluorescent* was studied, and it diversification observed on relatively short time scales, in addition, some of the population branches were reported to engage in cooperative behaviors.

Evolutionary and ecological literature, while recognizing competition and the resulting negative frequency-dependent selection as an important mechanism of diversification, also suggests that facilitation[3] could provide an alternative path to speciation [10]. An example of facilitation where the presence of one species enhances the fitness of the other is cross-feeding [17, 43], a relatively common phenomenon in bacteria. For example, [49] describes evolutionary diversification that occurred primarily as a result of one species of *E. coli* having a facilitative effect on the second through introduction of additional resources into the environment. Mathematical models explaining the evolution of cross-feeding have been proposed

[3]i.e. species interactions that benefit at least one of the participants and cause harm to neither

in [14, 40].

In the present study we focus on mutualism, which is a special type of facilitation involving reciprocally positive interactions between two species, with benefits experienced by both species [7]. There are different kinds of mutualistic behavior, which, to different degrees, can be classified as cooperation. By-product mutualism implies that species A benefits by B's actions without altering the costs or benefits experienced by B (and vice versa), whereas pseudo-reciprocity and reciprocity involve costs associated with the cooperative behavior [9]. We have shown that mutualism of this kind can lead to speciation in the context of division of labor games. This result fits with other theories that have been proposed in the field.

In [16], cooperation is studied in the context of the snowdrift game [48]. In [16], there is a single trait evolving over time which represents the individuals' investments to the common good. The benefit that the individuals receive through their behaviors is a function of the sum of the investments to the common good. The individuals must also pay a cost associated with their own behavior. In this scenario there are parameter sets such that individuals branch into populations of cooperators and defectors. Similar to the model developed in [16], our model has individuals which contribute to the common good and when two of them interact, the sum of the common good is used to determine the benefit received by a participant. In addition, individuals must pay a cost associated with that behavior. In contrast to the model of [16], our model is inspired by a population of individuals which need to perform multiple sub-tasks of a complex task, rather than performing a single task (as in the snowdrift game). Therefore, the nature of the branching behavior is quite different from that found in [16]: instead of splitting into cooperators and defectors, both population branches continue to contribute to the common good, but each specializes on one of the two sub-tasks. Introducing defectors in our setting allows for further stratification of the population, where the two groups of specialists coexist with a group of defectors.

Another model of cooperation was explored in [50]. Similar to the present report, the authors use the division of labor setting and examine the scenario in which a population must perform two tasks to be successful. In their model, the traits that are evolved are the probability for choosing a specific task and the cost associated with that task. It is found that division of labor in a social population can lead to specialization. In this model, the individuals are forced to be specialists in each game, as they (probabilistically) pick one of the two tasks that they perform in each of the group interactions. The important difference in our setting is the possibility for individuals to be specialists or generalists: an individual may specialize, but it may also chose to not specialize and perform enough of each task to be independent; alternatively, an individual may perform or not perform any task while benefitting from other individual's efforts. The absence of restriction to specialization in our model defines the nature of the branching trajectories reported here.

Finally, there is a clear connection between our model and The Black Queen Hypothesis (BQH) proposed in [37]. The BQH specifies the conditions under which it is advantageous for an organism to stop performing a function. A loss of a functional gene may result in a loss of fitness. On the other hand, it is possible that the function was costly, and thus losing it may be associated with a gain, determined by the saving in energy or other resources. If the gain exceeds the loss, it is advantageous for the organism to lose the function. This situation can occur when the gene product is "leaky". The BQH predicts that the loss of a costly, leaky function can be selected for, and will continue on an individual level until an equilibrium is reached. At this stage, the population will consist of "beneficiaries" that have lost the function, and the "helpers" that continue to produce it as public goods. Our model is complementary to this hypothesis, in that it provides a mathematical basis for the BQH. It makes similar assumptions to the BQH: it is assumed that the organisms are in a situation in which resources produced by individuals are released into the environment to be used as a public good, and that by no longer performing a task, the individual will have an advantage. At the same time our framework is somewhat broader as it includes multiple mutations and the possibility of two tasks. Our model provides further support of the BQH by verifying that cooperation can lead to task specialization, but also complements it by offering the possibility that cooperation leading to speciation is what drove the gene loss and task specialization. For example, the cyanobacteria *Prochlorococcus* and *Synechococcus* are thought to share a common ancestor, and our model suggests that the loss of the defense gene *katG* in *Prochlorococcus* could have occurred due to part of the ancestral population maintaining this gene (evolving to become *Synechococcus*) while the rest of the population was able to disregard it (evolving to become *Prochlorococcus*) [37].

The present paper develops a theory of speciation in the presence of cooperation, but does not address the very origin and maintenance of cooperation. While a large amount of literature has been devoted to the question of the evolution of cooperation, see e.g. [1, 38], this is not the focus of the present paper. Instead, we assume that cooperative behavior is already in place (such as in the context of "leaky" products), and investigate consequences of that in terms of speciation and adaptation. We further explore the existence of cheaters and find that the evolutionary scenario found here is quite robust to the introduction of cheaters, and population specialization persists under a wide range of parameters. If, however, the rate of cheater mutation is higher than a threshold, the population of non-cheaters can be driven to extinction by the cheaters taking over. Alternatively, the presence of cheaters can prevent evolutionary branching from occurring, which results in a population that consists of a number of non-specialized non-cheaters maintaining a population of cheaters. Biologically this is reasonable. If specialized behavior is beneficial, then a small number of non-cooperative individuals will not change that. However, if there are enough cheaters, the non-cheaters will more commonly interact with cheaters, and will need to become more self sufficient. Finally, if the mutation rate from non-cheaters to cheaters is very high, then the non-cheaters may not be able to adapt quickly enough to the environment before the cheaters take over.

The approach used in this paper is appropriate for studying speciation, as it includes the possibility of a continuous change of a trait. Extensions would include more complicated population structures, such as spatial interactions, or interactions on networks. Further, more detailed studied of the cheater dynamics are necessary, including the question of " evolutionary suicide" [21, 22] which requires modifications to the model to account for changing population size. This is work in progress.

To conclude, it is our hope that the theory of diversification rooted in cooperation may inspire further ecological research to identify more evolutionary examples consistent with this speciation mechanism. In particular, it would be desirable to find evidence that cooperation over sufficiently long time can give rise to inheritable changes by evolutionary branching.

Chapter 4

Trait Dynamics for Evolutionary Systems: Dynamical Rules and Number of Interactions Impact Results

4.1 Introduction

Adaptive dynamics investigates how a trait value used to describe a behavior changes over time. The papers working out the theoretical basis for adaptive dynamics were introduced in the 1990's [36, 12, 19], and adaptive dynamics was more recently summarized in The Hitchhiker's Guide to Adaptive Dynamics [4], which covers single trait evolution for Adaptive Dynamics. In the theoretical derivations, a key concept is the fitness of each behavior within its environment. This fitness is calculated based on the interactions of the individual with all other members of the population (i.e., theoretically, in the framework of the ODEs, the number of interactions is infinite). An individual with a higher fitness will be able to create offspring, and one with lower fitness will see its behavior die out. In computer simulations, however, it may become computationally expensive to calculate the fitness of each individual based on its interactions with every other individual. Also, in order to provide a model that is realistic, one has to restrict the number of interactions that each individual takes part in during its life-span. Therefore, there is a certain discrepancy between computer simplifications using a limited number of interactions, and theory assuming an infinite number of interactions.

An example of such discrepancy is provided by a paper which discusses cooperation [16]. There, a computer simulation is used to illustrate how individuals will branch into cooperators and defectors in the snowdrift game (an idealized scenario often studied in game theory). In their model, the authors consider individuals with trait values that identify their cooperativeness. These trait values evolve over time based on their interactions with others. In the iterative steps of the simulation, two individuals are chosen: the individual to be replaced and its potential replacement. Each individual has their fitness calculated through a single interaction with another individual of the population. The fitness of the to-be-replaced individual is subtracted from the potential replacement's fitness. This difference is scaled by a number that ensures that all such differences are always less than one, and this scaled difference then serves as the probability of the replacement to take place. The question is whether this simplification in the simulation will still capture the behavior expected by the adaptive dynamics theory.

Another reasonable simplification (or rather, an alternative, but reasonable, modeling assumption) concerns with the exact way in which fitness differences translate into reproductive success. Rather than using the scaled difference as the probability of replacement, one could assume simply that the focal individual is replaced by a competitor if the competitor performs better. This simplification no longer requires the knowledge of the maximum or minimum value that the function takes, and so the simulation can be run without bounding the trait values.

As it turns out, both of these assumptions may lead to differences when compared to the adaptive dynamics, which we will investigate here. We will start with developing ordinary differential equations to describe the evolution of these simulations and discuss how, and under which circumstances they differ from the equations from the adaptive dynamics theory.

4.2 Equation for the Mean Trait Value Evolution

Following [36, 12, 19], we consider a population of individuals each characterized by a single trait value, x, which may be restricted to a collection of reasonable values. For example, if we are considering a population of deer and we wish to characterize antler size by a trait value, we would wish to restrict the trait value to only positive values. These dynamics can be simulated in a variety of ways, but for this paper, we will focus on a simulation in which there is a fixed population size of individuals, each with a single trait value. In this simulation, traits change through two mechanisms: replacement and mutation. Replacement occurs through a competition between two individuals, in which a payoff function determines how well an individual's trait does in the surrounding population. Namely, the payoff functions between the individuals are compared, and depending on the selection criterion for determining the winner of this interaction, one individual's trait value may be replaced by its competitor's trait value (or by its own). The trait value that is replacing the old trait value has a chance to mutate to a nearby value. For concreteness, we will be considering a mutation that is pulled from a normal distribution centered at the replacement's trait value. In this way, the population always remains the same size. We assume that the chance for a mutation to occur is sufficiently small that mutations are very rare, so that a new mutant value has a chance to either die out or invade the entire population before a new mutation is created. As a result of this, the population can be considered homogeneous (monomorphic) with respect to trait x.

We will start our derivation by examining the probability of a population expressing trait x at a time t. To understand the dynamics of the trait value x, we need to know which trait will replace another, and when this can happen. Consider what can cause the population to have trait x at time $t + \Delta t$. A population will have trait x at time $t + \Delta t$ under two circumstances.

- First, the population may have value y and become x through the creation of a mutant with trait x, which can invade the population of trait y. Denote the probability of being y at time t as $P(y, t)$. The probability of mutation during the infinitesimal time interval Δt is denoted by $\mu \Delta t$, and the probability distribution of different values of the mutant trait x produced by an individual with trait y is represented by $M(y, x)$, with $\int M(y, x)\, dy = 1$. In this study, we assume that $M(y, x)$ is a bell-like function of x concentrated around y, such that the probability of very large mutations is very small. We must also include the probability that a mutant with trait x is able to invade and become the dominant behavior, designated by $\phi(x, y)$. Therefore, the probability to create a successful mutant x during Δt is given by

  $$\mu M(y, x) \phi(x, y) P(y, t) \Delta t.$$

- The second way is for the population to be at value x and remain at x. The probability of being x at time t is $P(x, t)$. The probability of x not mutating away from x is

  $$(1 - \mu M(x, y) \phi(y, x) \Delta t) P(x, t),$$

  where $\phi(y, x)$ is the probability that y is able to supplant x.

We can sum these two parts to get the following Kolmogorov forward equation:

$$
\begin{aligned}
P(x,t+\Delta t) &= \int_y \left(\mu M(y,x)\phi(x,y)P(y,t)\Delta t + (1-\mu M(x,y)\phi(y,x)\Delta t)P(x,t)\right) dy \\
P(x,t+\Delta t) &= \mu \int_y \left(M(y,x)\phi(x,y)P(y,t) - M(x,y)\phi(y,x)P(x,t)\right) \Delta t dy + P(x,t) \\
P(x,t+\Delta t) - P(x,t) &= \mu \int_y \left(M(y,x)\phi(x,y)P(y,t) - M(x,y)\phi(y,x)P(x,t)\right) dy \\
\frac{P(x,t+\Delta t) - P(x,t)}{\Delta t} &= \mu \int_y \left(M(y,x)\phi(x,y)P(y,t) - M(x,y)\phi(y,x)P(x,t)\right) dy \\
\Rightarrow \frac{dP(x,t)}{dt} &= \mu \int_y \left(M(y,x)\phi(x,y)P(y,t) - M(x,y)\phi(y,x)P(x,t)\right) dy
\end{aligned}
$$

This equation is used to calculate the evolution of the mean of x as time changes. We proceed by noticing that

$$
\int_x x\frac{dP(x,t)}{dt}dx = \frac{d}{dt}\int_x xP(x,t) = \frac{d}{dt}\langle x\rangle,
$$

such that

$$
\frac{d\langle x\rangle}{dt} = \mu \int_x \int_y \left(M(y,x)\phi(x,y)P(y,t) - M(x,y)\phi(y,x)P(x,t)\right) xdydx. \tag{4.1}
$$

By multiplying by x, integrating over x, and renaming variables x and y in the second term, we obtain the equation for the mean trait value:

$$
\frac{d\langle x\rangle}{dt} = \mu \int\int M(y,x)\phi(x,y)P(y,t)(x-y)dxdy. \tag{4.2}
$$

To determine $\phi(x,y)$ in equation (4.2), we must understand under which conditions a mutant with trait x will replace a resident with trait y. As mentioned before, replacement is dependent on the comparison between the payoffs received by an individual with trait x and one with trait y. We will use two criteria to determine the replacement value here:

1. Deterministic criterion for replacement: the trait value with the higher payoff will be used as the replacement.

2. Probabilistic criterion for replacement: the probability of replacement is proportional to the difference of the two payoffs; this is the algorithm used by Doebeli and colleagues in [16].

The functional form of $\phi(x, y)$ depends on the model formulation. In what follows, we will examine 4 possibilities, which differ by (i) whether only one interaction is used to determine payoff or all possible interactions are used, and (ii) whether the replacement happens according to deterministic or probabilistic criterium. We start with two cases (deterministic and probabilistic) with all interactions, followed by two cases of a single interaction.

4.3 Average Over All Interactions

We will start by examining equation (4.2) when the entire environment is used to determine which individual wins in the competition for replacement. In particular, suppose there are a total of N individuals, i of which have trait x and the remaining $N-i$ have trait y. We would like to know the probability that individual with trait x replaces an individual with trait y. We will use $F(x, y)$ to denote the payoff of an individual with trait x when it interacts with another individual. We then only have $i-1$ individuals in the population with trait x left (since one has already been chosen for the competition). In this setting, the average payoff of an individual with trait x in this population is given by

$$\frac{1}{N}\left((i-1)F(x, x) + (N-i)F(x, y)\right),$$

and the expected payoff of individual with trait y is

$$\frac{1}{N}\left(iF(y,x)+(N-i-1)F(y,y)\right).$$

For the payoff of an individual with trait y, we also have that there are only $N-i-1$ individuals with trait y left to choose from, so $F(y,y)$ is multiplied by $N-i-1$. In the next two subsections we will examine the cases of the deterministic and probabilistic criteria.

4.3.1 Probability of Fixation: Deterministic Criterion

If we use the deterministic criterion, the replacement of individual with trait y by an individual with trait x occurs if

$$\frac{1}{N-2}((i-1)F(x,x)+(N-i-1)F(x,y)-(i-1)F(y,x)-(N-i-1)F(y,y))>0.$$

The left hand side of this inequality is the average value of the payoff function for an individual with trait x, less the average payoff of an individual in the environment (the environment in this equation is assumed to have trait value y). We are assuming the two competing individuals do not interact with each other, so there are $N-2$ individuals remaining to behave as cooperators. To determine the probability of fixation, we must start by examining the probability $p(i)$ that the number of individuals with behavior x grows from i individuals in the population to $i+1$:

$$p(i)=\begin{cases}\frac{1}{N-2}\Big((i-1)F(x,x)+(N-i-1)F(x,y) \\ \qquad -(i-1)F(y,x)-(N-i-1)F(y,y)\Big) & \text{if positive} \\ 0 & \text{otherwise}\end{cases} \tag{4.3}$$

Likewise, the probability that the number of individuals with behavior x drops from i to $i-1$, $q(i)$ is:

$$q(i) = \begin{cases} \frac{1}{N-2}\Big((i-1)F(y,x) + (N-i-1)F(y,y) \\ \qquad\qquad -(i-1)F(x,x) - (N-i-1)F(x,y)\Big) & \text{if positive} \\ 0 & \text{otherwise} \end{cases}$$

Recall that we are assuming that the probability of large mutations is very small, such that the trait of newly produced mutants is typically very close to the trait value of the resident population. Therefore we can expand the functions in Taylor series in terms of the trait value differences and only keep the first order terms. This allows us to write $p(i)$ as follows (similar calculations are performed for $q(i)$):

$$\begin{aligned} p(i) &= \frac{1}{N-2}\left((i-1)F(x,x) + (N-i-1)F(x,y) - (i-1)F(y,x) - (N-i-1)F(y,y)\right) \\ &= \frac{1}{N-2}\left((i-1)\left(F(x,x) - F(y,x)\right) + (N-i-1)\left(F(x,y) - F(y,y)\right)\right) \\ &\approx \frac{1}{N-2}\Bigg((i-1)\left(F(x,x) - F(x,x) - \frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=x}(y-x)\right) + \\ &\quad (N-i-1)\left(-F(y,y) + F(y,y) + \frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y)\right)\Bigg) \\ &= \frac{1}{N-2}\left((i-1)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=x}(x-y) + (N-i-1)\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y)\right) \\ &= \frac{1}{N-2}\left(\left((i-1)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=x} + (N-i-1)\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}\right)(x-y)\right) \\ &= \frac{1}{N-2}\left(\left((N-2)\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y} + i\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=x} - \frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}\right)\right)(x-y)\right) \\ &\approx \frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y) \end{aligned}$$

leaving us with

$$p(i) \approx \begin{cases} \frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y) & \text{if positive} \\ 0 & \text{otherwise} \end{cases}$$

$$q(i) \approx \begin{cases} \frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(y-x) & \text{if positive} \\ 0 & \text{otherwise.} \end{cases}$$

Notice that in this case, quantities $p(i)$ and $q(i)$ do not depend on i and they can not both be positive. Therefore, the probability of fixation is either zero (when $q(i) > 0$ and $p(i) = 0$) or one (when $p(i) > 0$ and $q(i) = 0$). Since in $p(i)$ and $q(i)$ the derivative is multiplied by $x-y$, $p(i)$ will be positive when x is larger than y and the derivative is positive; it will also be positive when x is smaller than y and the derivative is negative. Therefore, the probability of fixation is determined by the sign of

$$\frac{\partial F(z_1,x)}{\partial z_1}|_{z_1=x}(x-y).$$

If $x-y$ does not have the same sign as the derivative, then trait x will not be able to take hold of the population, and will die out. Therefore, the probability of invasion in equation (4.2) for this case is

$$\phi(x,y) = H\left(\frac{\partial F(z,y)}{\partial z}|_{z=y}(x-y)\right), \tag{4.4}$$

where H is the Heaviside function. We could now use this ϕ in the master equation, but first, we will find ϕ for the probabilistic criterion.

4.3.2 Probability of Fixation: Probabilistic Criterion

If we apply the probabilistic criterion for replacement in a scenario in which the entire population is considered, replacement of an individual with trait y by an individual with trait x will occur at a rate proportional to $iF(x,x)+(N-i)F(x,y)-iF(y,x)-(N-i)F(y,y)$. This is implemented in the simulation by replacement occurring when $(F(x,w)-F(y,z))/\alpha > r$, where α is a constant which keeps $0 < |(F(x,w)-F(y,z))|/\alpha < 1$ for all x, y, w, and z, and r is a random number uniformly distributed between zero and one. When we include all the individuals in the calculations, again with $i-1$ individuals with trait x and $N-i-1$ individuals with trait y, we adjust the constant α, so that

$$0 < |((i-1)F(x,w)-(N-i-1)F(y,z))|/\alpha < 1$$

for all i, x, y, w, and z. Then we have

$$p(i) = \begin{cases} \frac{1}{\alpha(N-2)}\Big((i-1)F(x,x)+(N-i-1)F(x,y) \\ \qquad\qquad -(i-1)F(y,x)-(N-i-1)F(y,y)\Big) & \text{if} > \text{r} \\ 0 & \text{otherwise} \end{cases}$$

We have a similar statement for $q(i)$:

$$q(i) = \begin{cases} \frac{1}{\alpha(N-2)}\Big((i-1)F(y,x)+(N-i-1)F(y,y) \\ \qquad\qquad -(i-1)F(x,x)-(N-i-1)F(x,y)\Big) & \text{if} > \text{r} \\ 0 & \text{otherwise} \end{cases}$$

Notice that the simplifications made for equation (4.3) also apply here, so we have

$$p(i) \approx \begin{cases} \frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}(x-y)/\alpha & \text{if} > \text{r} \\ 0 & \text{otherwise} \end{cases}$$

$$q(i) \approx \begin{cases} \frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}(y-x)/\alpha & \text{if} > \text{r} \\ 0 & \text{otherwise} \end{cases}$$

As before, only one of $p(i)$ and $q(i)$ can be positive, so again, the population with trait x will either take over or die out, such that equation (4.4) for the probability of fixation holds for this case as well.

4.3.3 ODE for the Mean Trait Value

For both criteria we have obtained the same function for the probability of fixation, so continuing the derivation of the master equation can be performed for both cases together. We will use equation (4.4) in the equation for the mean trait value, eq. (4.2). To proceed, we note that for any a and b, we can write

$$H(ab) = H(a)H(b) + H(-a)H(-b). \tag{4.5}$$

This expression is one if a and b are either both positive or both negative, and is zero otherwise. Using the notation

$$\Delta x = x - y,$$

we can write:

$$
\begin{aligned}
\frac{d\langle x\rangle}{dt} =&\mu\int\int P(y,t)H(\frac{\partial F(z,y)}{\partial z}|_{z=y}\Delta x)M(|\Delta x|)\Delta x dy d\Delta x \\
=&\mu\int P(y,t)H(\frac{\partial F(z,y)}{\partial z}|_{z=y})dy\int M(|\Delta x|)H(\Delta x)\Delta x d\Delta x \\
&-\mu\int P(y,t)H(-\frac{\partial F(z,y)}{\partial z}|_{z=y})dy\int M(|\Delta x|)H(-\Delta x)\Delta x d\Delta x.
\end{aligned}
\tag{4.6}
$$

We also have

$$
\int H(\Delta x)M(|\Delta x|)\Delta x d\Delta x = \frac{1}{2}\langle|\Delta x|\rangle, \int H(-\Delta x)M(|\Delta x|)\Delta x d\Delta x = -\frac{1}{2}\langle|\Delta x|\rangle \tag{4.7}
$$

where $\langle|\Delta x|\rangle$ denotes the mean of the quantity $|\Delta x|$ under the symmetric distribution $M(|\Delta x|)$. Therefore,

$$
\begin{aligned}
\frac{d\langle x\rangle}{dt} &= \frac{\mu\langle|\Delta x|\rangle}{2}\int P(y,t)\big(H(\frac{\partial F(z,y)}{\partial z}|_{z=y}) - H(-\frac{\partial F(z,y)}{\partial z}|_{z=y})\big)dy \\
&= \frac{\mu\langle|\Delta x|\rangle}{2}\int P(y,t)sign(\frac{\partial F(z,y)}{\partial z}|_{z=y})dy.
\end{aligned}
$$

This can be simplified further if we decouple the deterministic equation (for the 1st moment) from the rest of the equations (for higher moments). The simplest way to do this is by approximating the mean of a function by the function of the mean, as is done in the literature [12]. We then obtain,

$$
\frac{d\langle x\rangle}{dt} = \frac{\mu\langle|\Delta x|\rangle}{2}sign\left(\frac{\partial F(z,\langle x\rangle)}{\partial z}|_{z=\langle x\rangle}\right) \tag{4.8}
$$

The resulting equation for the mean trait evolution indicates that the trait increases at a constant rate if it is advantageous and decreases if it is disadvantageous as defined by the derivative of the payoff.

Both the deterministic and probabilistic criteria have the same probability of fixation function, so this analysis holds for both criteria. Regardless of the criteria used, if the entire population's trait values are taken into consideration when determining replacement, the trait will move in the direction of the first partial derivative of the payoff function, taken with respect to the first variable.

It is interesting to compare equation (4.8) with the equation derived in [12], which reads

$$\frac{d\langle x\rangle}{dt} = \frac{\mu\sigma^2}{2}\tilde{N}(\langle x\rangle)\frac{\partial F(z,\langle x\rangle)}{\partial z}|_{z=\langle x\rangle}, \tag{4.9}$$

where the dynamics of the expected trait are determined not only by the sign of the derivative, but also by its magnitude, and not by the mean mutation step, but by its variance σ^2. In addition, the steady state population size, $\tilde{N}(\langle x\rangle)$ also changes the rate of convergence. In our model, the population size remains constant, but in other models, there is variation in population size dependent on the trait value. The reason for the differences between the two mean trait evolution equations is the different assumptions made during the development of the equations. First, since we assume there is a fixed population size, the birth rate is the same as the death rate. In the derivation of equation (4.9), the birth rate and death rate are not assumed to be equal, and are instead based on the trait values. In particular, the per capita birth rate $b(x,y)$ and the per capita death rate $d(x,y)$ are defined as the rates for a mutant with trait x in a monomorphic population of trait y. Then the probability of escaping extinction of a mutant with trait x in a population of trait y is approximated by

$$\phi(x,y) \approx \begin{cases} 1-\frac{d(x,y)}{b(x,y)} & \text{if } b(x,y) > d(x,y), \\ 0 & \text{otherwise.} \end{cases}$$

This quantity is a continuous increasing function of Δx, which, when expanded in the Taylor series, gives an expression $s_x(y)\Delta x$, where $s_x(y)$ is the growth rate of mutant x in a population of y. As a result, it introduces an additional factor of Δx under the integral sign, which, after integration with $M(\Delta x)$, yields the variance σ^2 as opposed to just mean absolute increment, $\langle|\Delta x|\rangle$. It also results in the actual value of the payoff derivative, and not just its sign, influencing the speed of evolution. This should be compared to our case, where $\phi(x, y)$ is the (discontinuous at 0) Heaviside function, which cannot be expanded in Taylor series in terms of Δx, thus changing the derivation procedure and the result.

The intuitive reason for the difference between the two equations is a larger number of degrees of freedom in the birth-death process with a non-constant population. There, a new mutant expands (or goes extinct) probabilistically, even if it is advantageous. In our system, its fate is decided from the start, and only depends on whether it is advantageous or not.

4.4 Single Interaction

In section 4.3, we found that the two replacement criteria were equivalent when the average payoff with the entire population is used to determine replacement. This however is not the only situation of interest. In reality, interaction with the whole population does not necessarily take place in the life-time of an individual. Therefore, we will also explore the dynamics when fewer interactions contribute into the fitness calculations. In particular, we would like to study the case of a single interaction. As before, we will assume a population of N individuals, with i individuals with trait x, and $N-i$ with trait y.

4.4.1 Deterministic Criterion

We will again start with the deterministic criterion, but this time with only a single individual used to calculate the payoff. The replacement of individual with trait y by an individual with trait x will occur under several scenarios: if $F(x, x) - F(y, y) > 0$, $F(x, x) - F(y, x) > 0$, $F(x, y) - F(y, y) > 0$ or $F(x, y) - F(y, x) > 0$. This method does not take into consideration the magnitude of the difference; it only considers which payoff is larger. We can represent the probability of the population of individuals with trait x increasing from i to $i+1$ individuals in the following way

$$
\begin{aligned}
p(i) = {} & \frac{i(N-i)}{N(N-1)(N-2)(N-3)} \\
& \Big((i-1)(i-2)H(F(x,x)-F(y,x)) \\
& + (i-1)(N-i-1)H(F(x,x)-F(y,y)) \\
& + (i-1)(N-i-1)H(F(x,y)-F(y,x)) \\
& + (N-i-1)(N-i-2)H(F(x,y)-F(y,y)) \Big).
\end{aligned}
\tag{4.10}
$$

Similarly, we have $q(i)$, the probability that the population of individuals with trait x decreases from i to $i-1$ individuals:

$$
\begin{aligned}
q(i) = {} & \frac{i(N-i)}{N(N-1)(N-2)(N-3)} \\
& \Big((i-1)(i-2)H(F(y,x)-F(x,x)) \\
& + (i-1)(N-i-1)H(F(y,x)-F(x,y)) \\
& + (i-1)(N-i-1)H(F(y,y)-F(x,x)) \\
& + (N-i-1)(N-i-2)H(F(y,y)-F(x,y)) \Big).
\end{aligned}
\tag{4.11}
$$

The fraction in front of each term represents the probability that a particular difference will arise in the simulation. For example, the probability that x and y are paired with x and y respectively happens with probability $(i-1)(N-i-1)/(N-2)(N-3)$. Note that the numerator of this expression is not $i(N-i)$ because one individual with trait y and one with trait x have already been chosen to be part of the interaction as competitors. Each such probability has a denominator of $(N-2)(N-3)$, so we factored it out to the front. In front is also the probability that x and y are chosen to compete: $i(N-i)$. We would like to be able to simplify equation (4.10). Again, since mutations are assumed to be very close to the resident population's trait value, we will start by using Taylor expansions of the differences in the payoff functions:

$$\begin{aligned}F(x,x)-F(y,x) &\approx F(y,x)+\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(x-y)-F(y,x)\\ &=\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(x-y)\end{aligned}$$

$$\begin{aligned}F(x,y)-F(y,y) &\approx F(y,y)+\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y)-F(y,y)\\ &=\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y)\end{aligned}$$

$$\begin{aligned}F(x,x)-F(y,y) &\approx F(y,y)+\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y)+\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}(x-y)-F(y,y)\\ &=\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y)+\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}(x-y)\end{aligned}$$

$$\begin{aligned}F(x,y)-F(y,x) &\approx F(y,x)+\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(x-y)+\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x}(y-x)-F(y,x)\\ &=\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(x-y)-\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x}(x-y).\end{aligned}$$

Using these approximations, now we have

$$\begin{aligned}p(i)=\frac{i(N-i)}{N(N-1)(N-2)(N-3)}&\Bigg((i-1)(i-2)H\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(x-y)\right)\\ &+(i-1)(N-i-1)H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}+\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\right)(x-y)\right)\\ &+(i-1)(N-i-1)H\left(\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}-\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x}\right)(x-y)\right)\\ &+(N-i-1)(N-i-2)H\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=x}(x-y)\right)\Bigg),\end{aligned}$$

$$
\begin{aligned}
q(i) = \frac{i(N-i)}{N(N-1)(N-2)(N-3)} \Bigg(& (i-1)(i-2) H\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(y-x)\right) \\
& + (i-1)(N-i-1) H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y} + \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\right)(y-x)\right) \\
& + (i-1)(N-i-1) H\left(\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y} - \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x}\right)(y-x)\right) \\
& + (N-i-1)(N-i-2) H\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=x}(y-x)\right) \Bigg).
\end{aligned}
$$

For ease of notation we will use the following vector:

$$
\mathcal{H}^+ = \begin{bmatrix}
H\left(\frac{\partial F(z_1,x)}{\partial z_1}\big|_{z_1=y}(x-y)\right) \\
H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y} + \frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right)(x-y)\right) \\
H\left(\left(\frac{\partial F(z_1,x)}{\partial z_1}\big|_{z_1=y} - \frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=x}\right)(x-y)\right) \\
H\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=x}(x-y)\right)
\end{bmatrix}
$$

$$
\mathcal{H}^- = \begin{bmatrix}
H\left(\frac{\partial F(z_1,x)}{\partial z_1}\big|_{z_1=y}(y-x)\right) \\
H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y} + \frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right)(y-x)\right) \\
H\left(\left(\frac{\partial F(z_1,x)}{\partial z_1}\big|_{z_1=y} - \frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=x}\right)(y-x)\right) \\
H\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=x}(y-x)\right)
\end{bmatrix}
$$

Note that $\mathcal{H}^+ + \mathcal{H}^-$ is a vector of all ones. This notation allows us to formally write the probabilities $p(i)$ and $q(i)$ as products between row vectors and column vectors:

$$
\begin{aligned}
p(i) = {} & \frac{i(N-i)}{N(N-1)(N-2)(N-3)} \\
& \cdot \begin{bmatrix} (i-1)(i-2) & (i-1)(N-i-1) & (i-1)(N-i-1) & (N-i-1)(N-i-2) \end{bmatrix} \mathcal{H}^+ \\
q(i) = {} & \frac{i(N-i)}{N(N-1)(N-2)(N-3)} \\
& \cdot \begin{bmatrix} (i-1)(i-2) & (i-1)(N-i-1) & (i-1)(N-i-1) & (N-i-1)(N-i-2) \end{bmatrix} \mathcal{H}^-
\end{aligned}
$$

Each of the Heaviside functions that appear in vectors $\mathcal{H}^+$ and $\mathcal{H}^-$ are either zero or one, depending on the payoff functions. Therefore, we can exhaustively calculate all the possibilities for the probability of fixation.

Since there are 4 different Heaviside functions in $\mathcal{H}^+$, we could have up to 16 possible combinations, but some we can eliminate. Since x is assumed to be very close to y, $\frac{\partial F(z_1,x)}{\partial z_1}\big|_{z_1=y} \approx \frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=x}$. Also, if $\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y} + \frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}$ and $\frac{\partial F(z_1,x)}{\partial z_1}\big|_{z_1=y} - \frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=x}$ are both positive (negative), then $\frac{\partial F(z_1,x)}{\partial z_1}\big|_{z_1=y}$ and $\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=x}$ must be positive (negative) as well. Lastly, both $\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y} + \frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}$ and $\frac{\partial F(z_1,x)}{\partial z_1}\big|_{z_1=y} - \frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=x}$ have the same probability of occurring, $\frac{(i-1)(N-i-1)}{(N-2)(N-3)}$, so if they have opposite signs we only have to do one calculation to get two of the probabilities of fixation. So, we only need to consider the following possibilities for $\mathcal{H}^+$:

$$\mathcal{H}^+ = \begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix}^T, \ \mathcal{H}^+ = \begin{bmatrix}1 & 1 & 0 & 1\end{bmatrix}^T,$$
$$\mathcal{H}^+ = \begin{bmatrix}0 & 1 & 0 & 0\end{bmatrix}^T, \ \mathcal{H}^+ = \begin{bmatrix}0 & 0 & 0 & 0\end{bmatrix}^T.$$

We can now determine the probability of fixation for a given population size. For $\mathcal{H}^+ = \left[\begin{smallmatrix}1 & 1 & 1 & 1\end{smallmatrix}\right]$, we can see immediately that the probability of the number of mutants decreasing is zero, so the mutant will fixate. Similarly, if $\mathcal{H}^+$ is all zeros, then the probability of the number of mutants increasing is zero, so the mutant will not fixate. For $\mathcal{H}^+ = \left[\begin{smallmatrix}1 & 1 & 0 & 1\end{smallmatrix}\right]$, when there is only one mutant, the probability that the mutant dies is zero:

$$q(1) = \frac{N-1}{N(N-1)(N-2)(N-3)} \begin{bmatrix}0 & 0 & 0 & (N-2)(N-3)\end{bmatrix} \begin{bmatrix}0 \\ 0 \\ 1 \\ 0\end{bmatrix} = 0.$$

If we consider the dynamics of the population of mutants as a birth and death process with a finite number of states and two absorbing states: extinction ($i = 0$) and fixation ($i = N$), then the probability that the birth-death process is absorbed by $i = 0$ is zero, and therefore it must eventually be absorbed by $i = N$, which is fixation by the mutant. Similarly for the remaining combination, $\mathcal{H}^+ = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}$, if there is only one mutant then the probability of increasing is zero:

$$p(1) = \frac{N-1}{N(N-1)(N-2)(N-3)} \begin{bmatrix} 0 & 0 & 0 & (N-2)(N-3) \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = 0,$$

so the probability of fixation is zero. From our analysis, we can see that if the first entry in $\mathcal{H}^+$ is one, the probability of fixation by the mutant population is one, but if the first entry in $\mathcal{H}^+$ is zero, then there is no chance of fixation. The first entry of $\mathcal{H}^+$ is one when $\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=x}(x-y)$ is positive, so $\phi = H\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=x}(x-y)\right)$, which is the same as in section 4.3, and so the subsequent analysis will also be the same as in section 4.3. Therefore, the differential equation describing the evolution of the mean trait value is the same as equation (4.8):

$$\frac{d\langle x\rangle}{dt} \approx \frac{\mu\langle|\Delta x|\rangle}{2} sign\left(\frac{\partial F(z, \langle x\rangle)}{\partial z}\big|_{z=\langle x\rangle}\right)$$

4.4.2 Probabilistic Criterion

Now we will examine how the probabilistic criterion fares with a single interaction to determine the behavior of the trait values in the simulation. In this case, replacement of an individual with trait y by an individual with trait x also can occur in several scenarios, but the probability of replacement is proportional to the difference between the payoffs. This is achieved by finding an α such that

$$0 \leq |F(x,w) - F(y,z)| \leq 1$$

for any x, w, y, or z in the region that the trait values have been restricted to. Then, the probability of the population of individuals with trait x increasing from i to $i+1$ in a population of individuals with trait y is

$$\begin{aligned}
p(i) = \frac{i(N-i)}{N(N-1)(N-2)(N-3)}\Big(&(i-1)(i-2)H(F(x,x)-F(y,x))(F(x,x)-F(y,x))/\alpha \\
&+ (i-1)(N-i-1)H(F(x,x)-F(y,y))(F(x,x)-F(y,y))/\alpha \\
&+ (i-1)(N-i-1)H(F(x,y)-F(y,x))(F(x,y)-F(y,x))/\alpha \\
&+ (N-i-1)(N-i-2)H(F(x,y)-F(y,y))(F(x,y)-F(y,y))/\alpha\Big),
\end{aligned}$$

and the probability of this population decreasing from i to $i-1$ is

$$\begin{aligned}
q(i) = \frac{i(N-i)}{N(N-1)(N-2)(N-3)}\Big(&(i-1)(i-2)H(F(y,x)-F(x,x))(F(y,x)-F(x,x))/\alpha \\
&+ (i-1)(N-i-1)H(F(y,x)-F(x,y))(F(y,x)-F(x,y))/\alpha \\
&+ (i-1)(N-i-1)H(F(y,y)-F(x,x))(F(y,y)-F(x,x))/\alpha \\
&+ (N-i-1)(N-i-2)H(F(y,y)-F(x,y))(F(y,y)-F(x,y))/\alpha\Big).
\end{aligned}$$

As before, we will simplify using a Taylor expansion to get the following two statements:

$$
\begin{aligned}
p(i) = &\frac{i(N-i)}{\alpha N(N-1)(N-2)(N-2)}\\
&\Bigg((i-1)(i-2)H\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(x-y)\right)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(x-y)\\
&+(i-1)(N-i-1)H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}+\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\right)(x-y)\right)\\
&\quad\cdot\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}+\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\right)(x-y)\\
&+(i-1)(N-i-1)H\left(\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}-\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x}\right)(x-y)\right)\\
&\quad\cdot\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}-\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x}\right)(x-y)\\
&+(N-i-1)(N-i-2)H\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y)\right)\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y)\Bigg),\\
q(i) = &\frac{i(N-i)}{\alpha N(N-1)(N-2)(N-3)}\\
&\Bigg((i-1)(i-2)H\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(y-x)\right)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(y-x)\\
&+(i-1)(N-i-1)H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}+\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\right)(y-x)\right)\\
&\quad\cdot\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}+\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\right)(y-x)\\
&+(i-1)(N-i-1)H\left(\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}-\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x}\right)(y-x)\right)\\
&\quad\cdot\left(\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}-\frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x}\right)(y-x)\\
&+(N-i-1)(N-i-2)H\left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(y-x)\right)\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(y-x)\Bigg).
\end{aligned}
$$

As in the deterministic case, note that all but the last term in the probability of increasing or decreasing depend on $(i-1)$. When $i=1$, each of these terms will be zero, and the last term will be nonzero in either $p(1)$ or $q(1)$, but not both. If it is in $p(1)$, then the probability of decreasing from one mutant to none is zero, but if the term is present in $q(i)$, then the probability of increasing from one mutant to two is zero. So, each mutant produced either cannot die out, or it cannot increase in size. Therefore, the probability of fixation is one if

$\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}(x-y) > 0$ and zero otherwise. This leaves us with the same differential equation for describing the evolution of the trait values as for the deterministic method:

$$\frac{d\langle x\rangle}{dt} \approx \frac{\mu\langle|\Delta x|\rangle}{2} sign\left(\frac{\partial F(z,\langle x\rangle)}{\partial z}|_{z=\langle x\rangle}\right)$$

4.4.3 On The Importance of Time Scales for the Probabilistic Criterion

In the derivation of the previous equations, we made the assumption that fixation will complete before the next mutant arises. In the case of a single interaction when using the probabilistic criterion, we can see the importance clearly, as the magnitude of the partial derivatives impacts the rate of fixation. The reason for this will become clear if we investigate these probabilities of increase and decrease a bit further. For ease of notation, we denote

$$\begin{aligned}\mathcal{F}(y) &= \frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y} \approx \frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}\\ \mathcal{G}(y) &= \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x} \approx \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\end{aligned}$$

In the calculations performed above for the model with a single interaction, probabilistic criterion, three combinations of these functions appear within the Heaviside function in the expressions for $p(i)$ and $q(i)$: $\mathcal{F}(y), \mathcal{F}(y)+\mathcal{G}(y)$, and $\mathcal{F}(y)-\mathcal{G}(y)$. Note however that if both $\mathcal{F}(y)+\mathcal{G}(y)$ and $\mathcal{F}(y)-\mathcal{G}(y)$ are positive (negative), then $\mathcal{F}(y)$ must be positive (negative) as well. This leaves us with exactly 6 combinations of signs of the three quantities:

$$1.\begin{cases}\mathcal{F}(y)>0\\ \mathcal{F}(y)+\mathcal{G}(y)>0\\ \mathcal{F}(y)-\mathcal{G}(y)>0\end{cases}\quad 2.\begin{cases}\mathcal{F}(y)>0\\ \mathcal{F}(y)+\mathcal{G}(y)>0\\ \mathcal{F}(y)-\mathcal{G}(y)<0\end{cases}\quad 3.\begin{cases}\mathcal{F}(y)>0\\ \mathcal{F}(y)-\mathcal{G}(y)>0\\ \mathcal{F}(y)+\mathcal{G}(y)<0\end{cases}$$

$$4.\begin{cases}\mathcal{F}(y)<0\\ \mathcal{F}(y)+\mathcal{G}(y)>0\\ \mathcal{F}(y)-\mathcal{G}(y)<0\end{cases}\quad 5.\begin{cases}\mathcal{F}(y)<0\\ \mathcal{F}(y)-\mathcal{G}(y)>0\\ \mathcal{F}(y)+\mathcal{G}(y)<0\end{cases}\quad 6.\begin{cases}\mathcal{F}(y)<0\\ \mathcal{F}(y)-\mathcal{G}(y)<0\\ \mathcal{F}(y)+\mathcal{G}(y)<0\end{cases}$$

or equivalently

$$\begin{aligned}&1.\quad 0<|\mathcal{G}(y)|<\mathcal{F}(y),\quad 2.\quad 0<\mathcal{F}(y)<\mathcal{G}(y),\quad 3.\quad \mathcal{G}(y)<-\mathcal{F}(y)<0,\\ &4.\quad 0<-\mathcal{F}(y)<\mathcal{G}(y),\quad 5.\quad \mathcal{G}(y)<\mathcal{F}(y)<0,\quad 6.\quad 0<|\mathcal{G}(y)|<-\mathcal{F}(y)\end{aligned} \tag{4.12}$$

The probabilities $p(i)$ and $q(i)$ can be written concisely as

$$\begin{aligned}p(i)=&\frac{iN\Delta x}{\alpha N(N-1)(N-2)(N-3)}\\ &\Big[((i-1)(i-2)+(N-i-1)(N-i-2))H(\mathcal{F}\Delta x)\mathcal{F}\\ &+(i-1)(N-i-1)\Big(H((\mathcal{F}+\mathcal{G})\Delta x)(\mathcal{F}+\mathcal{G})+H((\mathcal{F}-\mathcal{G})\Delta x)(\mathcal{F}-\mathcal{G})\Big)\Big],\end{aligned} \tag{4.13}$$

$$\begin{aligned}q(i)=&\frac{-iN\Delta x}{\alpha N(N-1)(N-2)(N-3)}\\ &\Big[((i-1)(i-2)+(N-i-1)(N-i-2))H(-\mathcal{F}\Delta x)\mathcal{F}\\ &+(i-1)(N-i-1)\Big(H(-(\mathcal{F}+\mathcal{G})\Delta x)(\mathcal{F}+\mathcal{G})+H(-(\mathcal{F}-\mathcal{G})\Delta x)(\mathcal{F}-\mathcal{G})\Big)\Big].\end{aligned} \tag{4.14}$$

We will examine the ratio between equations (4.13) and (4.14). If the ratio is close to one, then the population has a similar chance to increase as it does to decrease, resulting in a very slow rate of fixation.

Let us denote $p^{k+}(i)$ and $q^{k+}(i)$ the probabilities to increase the population of mutants in case k, $k \in \{1, 2, 3, 4, 5, 6\}$ given that $\Delta x > 0$, that is, the mutant trait is larger than that of the residents. Similarly, we denote $p^{k-}(i)$ and $q^{k-}(i)$ these probabilities when $\Delta x < 0$. For each of the six cases, we assume that the appropriate inequalities in (4.12) hold. The following expressions can be derived for the ratios of probabilities of the six cases:

$$S_1: \quad \frac{q^{1+}(i)}{p^{1+}(i)} = \frac{q^{6-}(i)}{p^{6-}(i)} = 0, \tag{4.15a}$$

$$\begin{aligned} S_2: \quad \frac{q^{2+}(i)}{p^{2+}(i)} &= \frac{q^{5-}(i)}{p^{5-}(i)} \\ &= \frac{|\mathcal{F}-\mathcal{G}|(i-1)(N-i-1)}{((i-1)(i-2)+(N-i-1)(N-i-2))|\mathcal{F}| + (i-1)(N-i-1)|\mathcal{F}+\mathcal{G}|}, \end{aligned} \tag{4.15b}$$

$$\begin{aligned} S_3: \quad \frac{q^{3+}(i)}{p^{3+}(i)} &= \frac{q^{4-}(i)}{p^{4-}(i)} \\ &= \frac{|\mathcal{F}+\mathcal{G}|(i-1)(N-i-1)}{((i-1)(i-2)+(N-i-1)(N-i-2))|\mathcal{F}| + (i-1)(N-i-1)|\mathcal{F}-\mathcal{G}|}, \end{aligned} \tag{4.15c}$$

$$\begin{aligned} S_4: \quad \frac{q^{4+}(i)}{p^{4+}(i)} &= \frac{q^{3-}(i)}{p^{3-}(i)} \\ &= \left(\frac{((i-1)(i-2)+(N-i-1)(N-i-2))|\mathcal{F}| + (i-1)(N-i-1)|\mathcal{F}-\mathcal{G}|}{|\mathcal{F}+\mathcal{G}|(i-1)(N-i-1)}\right), \end{aligned} \tag{4.15d}$$

$$\begin{aligned} S_5: \quad \frac{q^{5+}(i)}{p^{5+}(i)} &= \frac{q^{2-}(i)}{p^{2-}(i)} \\ &= \left(\frac{((i-1)(i-2)+(N-i-1)(N-i-2))|\mathcal{F}| + (i-1)(N-i-1)|\mathcal{F}+\mathcal{G}|}{|\mathcal{F}-\mathcal{G}|(i-1)(N-i-1)}\right), \end{aligned} \tag{4.15e}$$

$$S_6: \quad \frac{q^{6+}(i)}{p^{6+}(i)} = \frac{q^{1-}(i)}{p^{1-}(i)} = (0)^{-1} \tag{4.15f}$$

(the last line simply means that $p^{6+}(i) = p^{1-}(i) = 0$). We notice the existence of equivalence between pairs of scenarios. For example, case 2 (when $\mathcal{F} > 0$, $\mathcal{F}+\mathcal{G} > 0$, $\mathcal{F}-\mathcal{G} < 0$) with $\Delta x > 0$ and case 5 (when $\mathcal{F} < 0$, $\mathcal{F}+\mathcal{G} < 0$, $\mathcal{F}-\mathcal{G} > 0$) with $\Delta x < 0$ are equivalent, as is evident by examining S_2 in equation (4.15). Because of such equivalences, there are altogether 6 distinct values of the $q(i)/p(i)$ ratios, even though there are 12 such ratios (six cases, each with the possibility of $\Delta x > 0$ and $\Delta x < 0$). Above we denoted the six distinct scenarios as $S_1, \ldots, S_6$. Note that the $q(i)/p(i)$ ratios for scenarios S_4 and S_5 are exactly the inverse of the $q(i)/p(i)$ ratios of scenarios S_3 and S_2, respectively. Further, for scenarios S_2 and S_5, we have from equation (4.12) that $\mathcal{G}/\mathcal{F} \geq 1$. When $\mathcal{G}/\mathcal{F} \to 1$, we have the ratios tend to 0 and ∞ for S_2 and S_5 respectively, and they both tend to 1 as $\mathcal{G}/\mathcal{F} \to \infty$. Similarly, for scenarios S_3 and S_4, we have from equation (4.12) that $(-\mathcal{G}/\mathcal{F}) \geq 1$. When $(-\mathcal{G}/\mathcal{F}) \to 1$, we have the ratios tend to 0 and ∞ for S_3 and S_4 respectively, and they both tend to 1 as $(-\mathcal{G}/\mathcal{F}) \to \infty$. These statements are summarized below:

$$S_2: \quad \frac{q^{1+}(i)}{p^{1+}(i)} = \frac{q^{4-}(i)}{p^{4-}(i)} \in [0,1), \tag{4.16}$$

$$S_3: \quad \frac{q^{2+}(i)}{p^{2+}(i)} = \frac{q^{3-}(i)}{p^{3-}(i)} \in [0,1), \tag{4.17}$$

$$S_4: \quad \frac{q^{3+}(i)}{p^{3+}(i)} = \frac{q^{2-}(i)}{p^{2-}(i)} \in (1,\infty), \tag{4.18}$$

$$S_5: \quad \frac{q^{4+}(i)}{p^{4+}(i)} = \frac{q^{1-}(i)}{p^{1-}(i)} \in (1,\infty). \tag{4.19}$$

So we can see that for four of the six cases, if $|\mathcal{G}/\mathcal{F}|$ is very large, then the ratio between $p(i)$ and $q(i)$ become very close to 1. This means that fixation will take a very long time, although it will happen. Therefore, if $|\mathcal{G}/\mathcal{F}|$ is large, we must choose smaller mutation rates to keep our assumptions of monomorphic populations.

In figure 4.1 we demonstrate these points by using the payoff function

$$P(x,y) = b_1(x+y) + b_2(x+y)^2 - c_1 x - c_2 x^2 + d_1 y. \tag{4.20}$$

For the specific parameter values used i the figure, we have $|\mathcal{G}/\mathcal{F}| = (b_1 + 2b_2 + d_1)/(b_1 + 2b_2 - c_1 - 2c_2) = 4812$. The figure shows the difference in convergence rates due to the noise introduced by the addition of multiple mutations present at once in the probabilistic case.

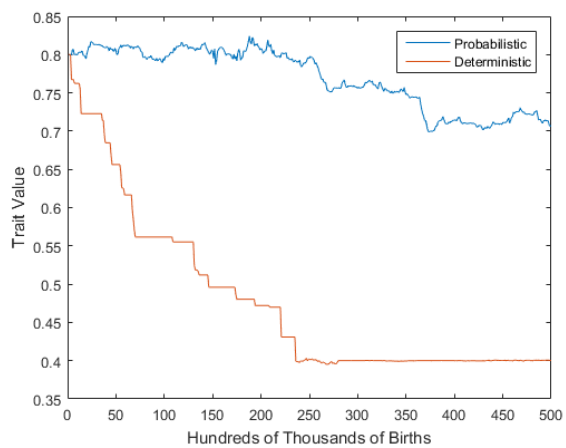


Figure 4.1: Comparison of convergence rates for the probabilistic criterion and the deterministic criterion. For the same mutation rate, the deterministic criterion converges quickly, but the probabilistic does not. In addition, the deterministic criterion consistently converges towards the steady state at .4, while the probabilistic is susceptible to mutations away from the steady state. The simulation using the probabilistic criterion may not match the ODE solution in this case. With such a slow rate of fixation, multiple mutations have the opportunity to appear, disrupting the expected behavior. By increasing d_1, the fixation rate will continue to slow when using the probabilistic criterion, but the simulation using the deterministic criterion will not. The payoff function is given by equation (4.20), with $b_1 = 6$, $b_2 = -1.4$, $c_1 = 4.56$, $c_2 = -1$ and $d_1 = 5000$ and mutation rate $\mu = .001$.

Since the deterministic method does not depend on the magnitude of the derivatives, but only on the sign of the derivatives or the sign of the sum or difference of the derivatives, it is not as susceptible to the derivative of the payoff function with respect to the environment. We will illustrate this difference by performing the same calculations, but for the deterministic, single interaction case. From previous calculations, we have:

$$\begin{aligned}
p(i) &= \frac{i(N-i)}{N(N-1)(N-2)(N-3)} \\
&\quad \cdot \begin{bmatrix} (i-1)(i-2) & (i-1)(N-i-1) & (i-1)(N-i-1) & (N-i-1)(N-i-2) \end{bmatrix} \mathcal{H}^{+} \\
q(i) &= \frac{i(N-i)}{N(N-1)(N-2)(N-3)} \\
&\quad \cdot \begin{bmatrix} (i-1)(i-2) & (i-1)(N-i-1) & (i-1)(N-i-1) & (N-i-1)(N-i-2) \end{bmatrix} \mathcal{H}^{-}.
\end{aligned}$$

Recall that $\mathcal{H}^{+}$ and $\mathcal{H}^{-}$ are each vectors of zeros and ones, and their sum is a vector of all ones. We have the same six cases as before, however, now cases two and three result in the same values of $p(i)$ and $q(i)$ and cases four and five also have the save values for $p(i)$ and $q(i)$. Since these values do not depend on the magnitude of the functions, we can exhaustively calculate these values:

1. $p(i) = \frac{i(N-i)}{N(N-1)(N-2)(N-3)}((i-1)(i-2) + 2(i-1)(N-i-1) + (N-i-1)(N-i-2))$,
   $q(i) = 0$.

2.,3. $p(i) = \frac{i(N-i)}{N(N-1)(N-2)(N-3)}((i-1)(i-2) + (i-1)(N-i-1) + (N-i-1)(N-i-2))$,
   $q(i) = \frac{i(N-i)}{N(N-1)(N-2)(N-3)}((i-1)(N-i-1))$

4.,5. $p(i) = \frac{i(N-i)}{N(N-1)(N-2)(N-3)}((i-1)(N-i-1))$,
   $q(i) = \frac{i(N-i)}{N(N-1)(N-2)(N-3)}((i-1)(i-2) + (i-1)(N-i-1) + (N-i-1)(N-i-2))$

6. $p(i) = 0$,
   $q(i) = \frac{i(N-i)}{N(N-1)(N-2)(N-3)}((i-1)(i-2) + 2(i-1)(N-i-1) + (N-i-1)(N-i-2))$

So, we can see that the ratio between $p(i)$ and $q(i)$ do not depend on the magnitude of the functions, only on their signs. In cases two through five, there is a positive chance to either increase or decrease in the number of mutants, so the time to fixation will be longer in those cases than in cases one and six, in which the probability to decrease or increase, respectively, is zero. In case two/three, we have

$$\begin{aligned} q(i)/p(i) &= \frac{(i-1)(i-2) + (i-1)(N-i-1) + (N-i-1)(N-i-2)}{(i-1)(N-i-1)} \\ &= 1 + \frac{(i-1)(i-2) + (N-i-1)(N-i-2)}{(i-1)(N-i-1)} > 1. \end{aligned}$$

In case four/five,

$$q(i)/p(i) = \frac{(i-1)(N-i-1)}{(i-1)(i-2) + (i-1)(N-i-1) + (N-i-1)(N-i-2)} < 1.$$

In both of these cases though, the value is determined by the number of mutants in the population, and so the payoff function only determines the case between one through six. So, once an appropriate mutation rate for the single interaction, deterministic criterion is found, it should provide consistent results regardless of the functional form of the payoff function. For the single interaction, probabilistic criterion, however, the payoff function will determine which mutation rates are appropriate.

4.5 Comparison Between Different Models

We will now examine the differences between the different models. We have seen that regardless of the number of interactions, both the probabilistic model and the deterministic model should behave the same. To illustrate the consistency between our derived ordinary differential equation and the simulation, we will consider a payoff function from [16]:

$$F(x, y) = b_1(x + y) + b_2(x + y)^2 - (c_1 x + c_2 x^2), \tag{4.21}$$

with $b_1 = 6$, $b_2 = -1.4$, $c_1 = 4.56$, $c_2 = -1$. From the ODE, we expect that the steady states will occur at y when

$$\frac{\partial F(z_1, y)}{\partial z_1}\Big|_{z_1=y} = 0.$$

In this particular example, this occurs at $y = .4$. The stability of this steady state is determined by the derivative

$$\frac{\partial}{\partial y}\left(\frac{\partial F(z_1, y)}{\partial z_1}\Big|_{z_1=y}\right),$$

which is -3.6 at the steady state $y = .4$, indicating that the steady state is in fact stable.

More information about the nature of the steady state can be obtained by using the formalism of adaptive dynamics. T determine whether the singular strategy (the steady state of the ODE) that we found above is an evolutionarily stable strategy, we examine the second derivative,

$$\frac{\partial^2 F(z_1, y)}{\partial z_1^2}\Big|_{z_1=y}.$$

At $y = .4$, this is $-.8 < 0$, so this is an evolutionarily stable strategy, indicating that in the simulation, once the trait value converges to .4, the trait value is expected to remain at that value (as opposed to undergoing branching, which would occur if this second derivative was positive).

Figure 4.2 shows the evolution of trait values when the average payoff over all individuals is

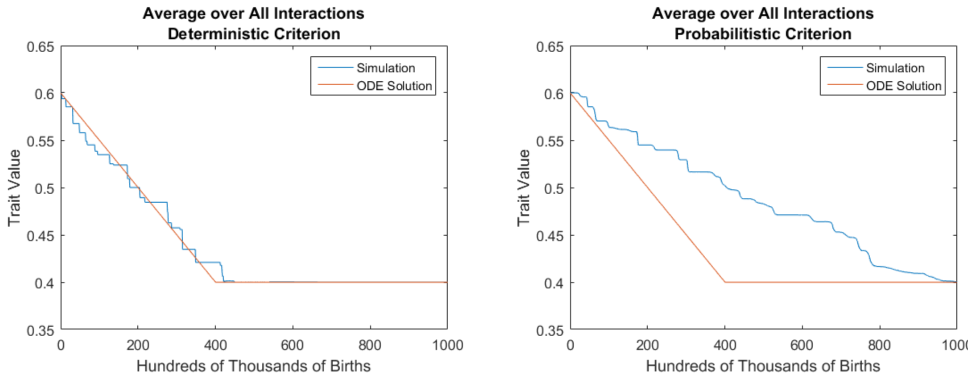


Figure 4.2: Plots of the average trait value in simulations of 100 individuals. The chance of mutation was 10^{-6} with a standard deviation of the mutant from the parent of .01. The average of the payoff with the environment was used for both images. Left: the deterministic criterion was used to determine replacement. Right: replacement was performed using the probabilistic criteria. The simulation using the probabilistic criterion is slow to switch to a mutant's behavior, so doesn't match the ODE time scale.

used for both the deterministic and probabilistic methods. In figure 4.2 (left), we can see that the deterministic criterion converges quickly for a relatively high rate of mutation. However, from figure 4.2 (right) we can also see that the probabilistic criterion does not converge at the rate predicted by the ODE. This is because mutations can arise before a new behavior has completely taken over a population. In the deterministic case, replacement occurs for differences greater than zero, but for the probabilistic case, replacement only occurs when the difference is greater than a uniformly distributed random number, slowing fixation. We can see improvement in the consistency between the simulation and differential equation solution in figure 4.3, which uses a smaller mutation rate.

We see similar consistency between the ordinary differential equation and the simulation when there is only a single interaction, as shown in figure 4.4. In the single interaction case, a smaller mutation rate is needed to see consistency between the ordinary differential equation and the simulation. This is because when only a single interaction is used, the time to fixation will be longer, meaning that the time between mutations must be longer to allow for the additional time for a mutant to fixate. This slower rate of fixation is due to the

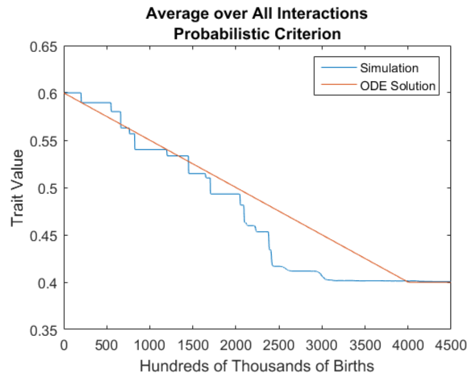


Figure 4.3: Plot of the average trait value of a simulation with 100 individuals. The chance of mutation was 10^{-7} (ten times smaller than in figure 4.2) with a standard deviation of the mutant from the parent of .01. The average of the payoff with the environment was used along with the probabilistic criterion. The rate of mutation is small enough here that the rate of convergence more closely matches the ODE time scale.

possibility that both $p(i)$ and $q(i)$ can be positive when a single interaction is used. This can cause the number of mutants to decrease, even for a beneficial mutant. For the case when the average over all interactions is used, either $q(i) = 0$ for all i or $p(i) = 0$ for all i, given a mutant trait value and environment trait value. Therefore, a beneficial mutation will only become more populous, and never decrease in number.

So, in the case in which there is a monomorphic population of individuals, the simulations will converge to the same trait values. Furthermore, we have shown that although we use the approximation

$$\int P(y,t) sign \left(\frac{\partial F(z,y)}{\partial z} |_{z=y} \right) dy = sign \left(\frac{\partial P(z, \langle x \rangle)}{\partial z} |_{z=\langle x \rangle} \right)$$

in the derivation of the ordinary differential equation, our simulations still closely match the solutions of the ordinary differential equation.

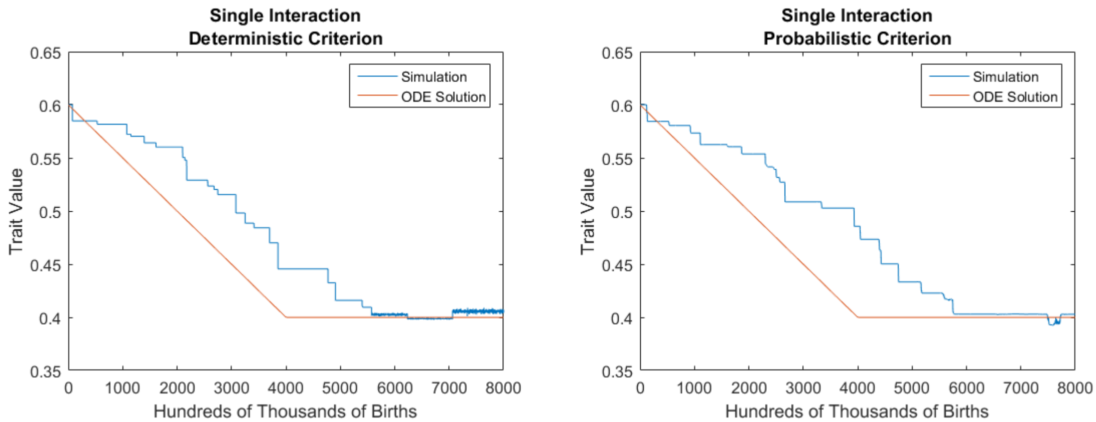


Figure 4.4: Plots of the average trait value in simulations of 100 individuals. The chance of mutation was 10^{-7} with a standard deviation of the mutant from the parent of .01. A single other individual was used to determine the payoff for both images. Left: replacement was performed using the deterministic criteria. Right: replacement was performed using the probabilistic criteria. Both simulations converge more slowly than the ODE, indicating that the rate of mutation is too high for the assumption in the derivation of the ODE that only one mutant trait is present at a time.

4.6 Including a Constant Cheater Population

We will now look at a case where two distinct trait values are present in the population. Rather that examine how two traits evolve when interacting, we will instead consider a simplification, in which one of the trait values will be held fixed, while the other evolves with time. This case is inspired by the model of cooperation and defection, where populations of cooperators and "cheaters" coexist as in chapter 3. Cheaters here are defined as individuals that do not produce the necessary products, but they can benefit from the surrounding population[1]. We will fix the number of cheaters to be constant; in the simulations this is done by preventing cheaters from being chosen as either individuals to be replaced or individuals to replace others. Since the evolving trait will still display monomorphic behavior, we will use the same mean trait evolution equation, eq. (4.2), as in previous sections. As before, to determine whether a given individual is replaced with the progeny of another individual, the

[1]In the somewhat artificial, fixed population scenario, cheaters can exist at any fixed population, but in a scenario of chapter 3, in which the population of cheaters does vary, their growth is dependent on the cooperators.

payoffs from interactions will be calculated and compared. Now, more scenarios are possible, as one could be paired with a cooperator or with a cheater, leading to different payoffs.

We will consider a population with i mutants, n residents and c cheaters, with $i+n+c=N$, the total constant population size of non-cheaters and cheaters. Cheaters are subject to the same payoff function as the non-cheaters, but cheaters will always have the trait value 0, which does not evolve over time.

4.6.1 Probability of Fixation: Average Over All Interactions

First, we will start by considering the case when the average payoff over all interactions is used. Consider a mutant population of i individuals with trait value x in an environment with n individuals with trait y and c cheaters. Using the deterministic replacement criterion, an individual with trait x will replace one with trait y with probability

$$\begin{aligned}
p(i) = & \frac{in}{(N-c)(N-c-1)(N-2)} \\
& \big((i-1)F(x,x) + (n-1)F(x,y) + cF(x,0) \\
& - ((i-1)F(y,x) + (n-1)F(y,y) + cF(y,0))\big) \\
\approx & \frac{in}{(N-c)(N-c-1)(N-2)} \\
& \left((i-1)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}(x-y) + (n-1)\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}(x-y)\right. \\
& \left. + c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}(x-y)\right) \\
\approx & \frac{in}{(N-c)(N-c-1)(N-2)}\left((n+i-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y} + c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)(x-y) \\
\approx & \frac{in}{(N-c)(N-c-1)(N-2)}\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y} + c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)(x-y)
\end{aligned}$$

We are only permitting individuals with trait x or y to be the competing individuals, so the chance that an individual with trait x competes with an individual with trait y is

$in/((N-c)(N-c-1))$. Then the individual with trait x interacts with $(i-1)$ individuals with trait x, with $(n-1)$ individuals with trait y, and with cheater individuals c times, for a total of $N-2$ interactions. The probability that an individual with trait y will replace one with trait x is similarly

$$q(i) \approx \frac{in}{(N-c)(N-c-1)}$$
$$\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y} + c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)(y-x).$$

If, on the other hand, we use the probabilistic criterion, we must define an α such that

$$\Big|(i-1)F(x,x)+(n-1)F(x,y)+cF(x,0)$$
$$-(i-1)F(y,x)+(n-1)F(y,y)+cF(y,0)\Big| < \alpha$$

for all i, x, and y in a closed region we define. Then the probability of increasing from i to $i+1$ mutants is

$$p(i) \approx \frac{in}{\alpha(N-c)(N-c-1)}$$
$$\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y} + c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)(x-y).$$

The probability of decreasing from i to $i-1$ mutants is

$$q(i) \approx \frac{in}{\alpha(N-c)(N-c-1)}$$
$$\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y} + c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)(y-x).$$

As in the previous calculations, we can see that for both deterministic and probabilistic criteria, either $p(i)$ or $q(i)$ will be positive for all i for a fixed set of parameters. Then, the probability of fixation is either one or zero, and is dependent on the sign of

$$\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}+c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)(x-y).$$

As a consequence, the probability of fixation for both the deterministic and probabilistic case when we average the payoff functions over all interactions is

$$\phi = H\left(\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}+c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)(x-y)\right)$$

Note that if there are no cheaters, this statement reduces down to equation (4.4).

With the probability of fixation, we can now determine the mean trait evolution equation. Using equations (4.5) and (4.7), we obtain,

$$\begin{aligned}
\frac{d\langle x\rangle}{dt} &= \frac{\mu\langle|\Delta x|\rangle}{2}\Bigg(\int P(y,t)\Bigg(H\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}+c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)\\
&\quad - H\left(-\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}+c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)\right)\Bigg)dy\\
&= \frac{\mu\langle|\Delta x|\rangle}{2}\Bigg(\int P(y,t)sign\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}+c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)dy\Bigg)\\
&\approx \frac{\mu\langle|\Delta x|\rangle}{2}sign\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}+c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right)
\end{aligned}$$

Since both the deterministic and probabilistic criteria have the same probability of fixation function, this analysis holds for both criteria. So, regardless of the criteria used, if the entire population's trait values are taken into consideration when determining replacement, the trait will evolve as predicted by the following mean trait evolution equation:

$$\frac{d\langle x\rangle}{dt} \approx \frac{\mu\langle\Delta x\rangle}{2}sign\left((N-c-2)\frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y}+c\frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y}\right) \qquad (4.22)$$

4.6.2 Single Interaction: Deterministic Criterion

Next, we examine the single interaction, deterministic case, where we must again determine the probability of fixation. The following equations describe how often the population of mutants grows ($p(i)$) and decays ($q(i)$) when there are i mutants:

$$\begin{aligned}
p(i) = {} & \frac{in}{(N-c)(N-c-1)(N-2)(N-3)} \\
& \cdot \Big((i-1)(i-2)H(F(x,x)-F(y,x)) + (i-1)(n-1)H(F(x,x)-F(y,y)) \\
& + (i-1)cH(F(x,x)-F(y,0)) + (n-1)(n-2)H(F(x,y)-F(y,y)) \\
& + (i-1)(n-1)H(F(x,y)-F(y,x)) + (n-1)cH(F(x,y)-F(y,0)) \\
& + (n-1)cH(F(x,0)-F(y,y)) + (i-1)cH(F(x,0)-F(y,x)) \\
& + c(c-1)H(F(x,0)-F(y,0))\Big),
\end{aligned} \tag{4.23}$$

$$\begin{aligned}
q(i) = {} & \frac{in}{(N-c)(N-c-1)(N-2)(N-3)} \\
& \cdot \Big((i-1)(i-2)H(F(y,x)-F(x,x)) + (i-1)(n-1)H(F(y,x)-F(x,y)) \\
& + (i-1)cH(F(y,x)-F(x,0)) + (n-1)(n-2)H(F(y,y)-F(x,y)) \\
& + (i-1)(n-1)H(F(y,y)-F(x,x)) + (n-1)cH(F(y,y)-F(x,0)) \\
& + (n-1)cH(F(y,0)-F(x,y)) + (i-1)cH(F(y,0)-F(x,x)) \\
& + c(c-1)H(F(y,0)-F(x,0))\Big)
\end{aligned} \tag{4.24}$$

The multipliers in front of each Heaviside function correspond to the probability of that interaction occurring. For example, the probability that x and y are competing and x is paired with x and y is also paired with x is $(in)(i-1)(i-2)/((N-c)(N-c-1)(N-2)(N-3))$. We are requiring the competitors to be non-cheaters, so there are only $N-c$ options for the first competitor and $N-c-1$ for the second. The assisting individuals are not restricted in the same way, so there are $(N-2)$ and $(N-3)$ choices for the cooperators. The term

$in/((N-c)(N-c-1)(N-2)(N-3))$ will be part of each statement, so it is factored out to the front.

As before, since x and y are very close, the differences of the payoff function can be approximated by derivatives. In addition, the differences $F(x,0)-F(y,x)$, $F(x,0)-F(y,y)$, $F(y,0)-F(x,x)$ and $F(y,0)-F(x,y)$ should have the same sign so long as $F(x,0)-F(x,x)$ is not close to zero:

$$\begin{aligned}
F(x,0)-F(y,x) &= F(x,0)-F(x,x)-\frac{\partial F(z,x)}{\partial z}|_{z=x}(y-x)\\
F(x,0)-F(y,y) &= F(x,0)-F(x,x)-\frac{\partial F(z,x)}{\partial z}|_{z=x}-\frac{\partial F(x,z)}{\partial z}|_{z=x}(y-x)\\
F(y,0)-F(x,x) &= F(x,0)-F(x,x)+\frac{\partial F(z,0)}{\partial z}|_{z=y}(x-y)\\
F(y,0)-F(x,y) &= F(x,0)-F(x,x)+\frac{\partial F(z,0)}{\partial z}|_{z=y}(x-y)-\frac{\partial F(x,z)}{\partial z}|_{z=x}(y-x)
\end{aligned}$$

This means that for the $c(n-1)$ and $c(i-1)$ term, only one of the two Heaviside functions is one. Utilizing these approximations, we can simplify both p and q:

$$\begin{aligned}
p(i) \approx{}& \frac{in}{(N-c)(N-c-1)(N-2)(N-3)}\\
&\cdot\Bigg((i-1)(i-2)H\left(\frac{\partial F(z_1,x)}{\partial z_1}|_{z_1=x}(x-y)\right)\\
&+(i-1)(n-1)H\left((\frac{\partial F(z_1,x)}{\partial z_1}|_{z_1=x}+\frac{\partial F(x,z_2)}{\partial z_2}|_{z_2=x})(x-y)\right)\\
&+(n-1)(n-2)H\left(\frac{\partial F(z_1,y)}{\partial z_1}|_{z_1=x}(x-y)\right)\\
&+(i-1)(n-1)H\left((\frac{\partial F(z_1,y)}{\partial z_1}|_{z_1=x}-\frac{\partial F(x,z_2)}{\partial z_2}|_{z_2=y})(x-y)\right)\\
&+c(c-1)H\left(\frac{\partial F(z_1,0)}{\partial z_1}|_{z_1=x}(x-y)\right)+(i-1)c+(n-1)c\Bigg),
\end{aligned} \tag{4.25}$$

$$
\begin{aligned}
q(i) \approx{}& \frac{in}{(N-c)(N-c-1)(N-2)(N-3)} \\
&\cdot \Bigg((i-1)(i-2) H \left(\frac{\partial F(z_1,x)}{\partial z_1}|_{z_1=x} (y-x) \right) \\
&+ (i-1)(n-1) H \left(\left(\frac{\partial F(z_1,x)}{\partial z_1}|_{z_1=x} + \frac{\partial F(x,z_2)}{\partial z_2}|_{z_2=x}\right)(y-x) \right) \\
&+ (n-1)(n-2) H \left(\frac{\partial F(z_1,y)}{\partial z_1}|_{z_1=x} (y-x) \right) \\
&+ (i-1)(n-1) H \left(\left(\frac{\partial F(z_1,y)}{\partial z_1}|_{z_1=x} - \frac{\partial F(x,z_2)}{\partial z_2}|_{z_2=y}\right)(y-x) \right) \\
&+ c(c-1) H \left(\frac{\partial F(z_1,0)}{\partial z_1}|_{z_1=x} (y-x) \right) + (i-1)c + (n-1)c \Bigg).
\end{aligned}
\tag{4.26}
$$

As before, we will use condensed notation

$$
\begin{aligned}
\mathcal{F}0(y) &= \frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y} \approx \frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=x} \\
\mathcal{F}(y) &= \frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y} \approx \frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y} \\
\mathcal{G}(y) &= \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x} \approx \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}
\end{aligned}
$$

There are four combinations of these functions present in the Heaviside functions: $\mathcal{F}0$, $\mathcal{F}$, $\mathcal{F}+\mathcal{G}$ and $\mathcal{F}-\mathcal{G}$. As before, if $\mathcal{F}+\mathcal{G}$ is positive (negative) and $\mathcal{F}-\mathcal{G}$ is positive (negative), then $\mathcal{F}$ is also positive (negative). In addition, since we are only considering the Heaviside function of these functions, from equations (4.25-4.26) we see that $H\left((\mathcal{F}+\mathcal{G})(x-y)\right)$ and $H\left((\mathcal{F}-\mathcal{G})(x-y)\right)$ are both multiplied by $(i-1)(n-1)$, so they each have the same impact on the behavior, and therefore we can consider the cases where $\mathcal{F}+\mathcal{G}$ and $\mathcal{F}-\mathcal{G}$ are opposite signs as a single case. So, although there are sixteen combinations for the Heaviside functions, by taking into account the restrictions due to the signs of the functions, there are only eight unique cases.

We can summarize the possibilities of the signs of these functions as:

$$
\begin{array}{llllllll}
1. \begin{cases} \mathcal{F}0 > 0 \\ \mathcal{F}+\mathcal{G} > 0 \\ \mathcal{F}-\mathcal{G} > 0 \\ \mathcal{F} > 0 \end{cases} & 2. \begin{cases} \mathcal{F}0 < 0 \\ \mathcal{F}+\mathcal{G} > 0 \\ \mathcal{F}-\mathcal{G} > 0 \\ \mathcal{F} > 0 \end{cases} & 3. \begin{cases} \mathcal{F}0 > 0 \\ \mathcal{F}+\mathcal{G} < 0 \\ \mathcal{F}-\mathcal{G} > 0 \\ \mathcal{F} > 0 \end{cases} \text{ and } \begin{cases} \mathcal{F}0 > 0 \\ \mathcal{F}+\mathcal{G} > 0 \\ \mathcal{F}-\mathcal{G} < 0 \\ \mathcal{F} > 0 \end{cases} \\
4. \begin{cases} \mathcal{F}0 < 0 \\ \mathcal{F}+\mathcal{G} < 0 \\ \mathcal{F}-\mathcal{G} > 0 \\ \mathcal{F} > 0 \end{cases} \text{ and } \begin{cases} \mathcal{F}0 < 0 \\ \mathcal{F}+\mathcal{G} > 0 \\ \mathcal{F}-\mathcal{G} < 0 \\ \mathcal{F} > 0 \end{cases} & 5. \begin{cases} \mathcal{F}0 > 0 \\ \mathcal{F}+\mathcal{G} < 0 \\ \mathcal{F}-\mathcal{G} > 0 \\ \mathcal{F} < 0 \end{cases} \text{ and } \begin{cases} \mathcal{F}0 > 0 \\ \mathcal{F}+\mathcal{G} > 0 \\ \mathcal{F}-\mathcal{G} < 0 \\ \mathcal{F} < 0 \end{cases} \\
6. \begin{cases} \mathcal{F}0 < 0 \\ \mathcal{F}+\mathcal{G} < 0 \\ \mathcal{F}-\mathcal{G} > 0 \\ \mathcal{F} < 0 \end{cases} \text{ and } \begin{cases} \mathcal{F}0 < 0 \\ \mathcal{F}+\mathcal{G} > 0 \\ \mathcal{F}-\mathcal{G} < 0 \\ \mathcal{F} < 0 \end{cases} & 7. \begin{cases} \mathcal{F}0 > 0 \\ \mathcal{F}+\mathcal{G} < 0 \\ \mathcal{F}-\mathcal{G} < 0 \\ \mathcal{F} < 0 \end{cases} & 8. \begin{cases} \mathcal{F}0 < 0 \\ \mathcal{F}+\mathcal{G} < 0 \\ \mathcal{F}-\mathcal{G} < 0 \\ \mathcal{F} < 0 \end{cases}
\end{array}
\tag{4.27}
$$

We will again denote $p^{k+}(i)$ and $q^{k+}(i)$ as the probabilities to increase and decrease the population of mutants in case k, $k \in \{1, \ldots, 8\}$, given that $\Delta x > 0$, that is, the mutant trait is larger than that of the resident. Similarly, we denote $p^{k-}(i)$ and $q^{k-}(i)$ as the probabilities when $\Delta x < 0$. Then, using $N = i + n + c$, we obtain:

$$S_1: \quad \frac{q^{1+}(i)}{p^{1+}(i)} = \frac{q^{8-}(i)}{p^{8-}(i)} = \frac{c(N-c-2)}{N^2 - N(c+5) + 3c + c^2 + 6}, \tag{4.28}$$

$$S_2: \quad \frac{q^{2+}(i)}{p^{2+}(i)} = \frac{q^{7-}(i)}{p^{7-}(i)} = \frac{c(N-2)}{(N-c-2)(N-3)} \tag{4.29}$$

$$\begin{aligned} S_3: \quad & \frac{q^{3+}(i)}{p^{3+}(i)} = \frac{q^{6-}(i)}{p^{6-}(i)} \\ & = \frac{(n-1)(i-1) + (i+n-2)c}{c^2 + (n-1)(i-1) + (i-1)(i-2) + (n-1)(n-2) + (i+n-2)c}, \end{aligned} \tag{4.30}$$

$$\begin{aligned} S_4: \quad & \frac{q^{4+}(i)}{p^{4+}(i)} = \frac{q^{5-}(i)}{p^{5-}(i)} \\ & = \frac{c^2 + (n-1)(i-1) + (i+n-2)c}{(n-1)(i-1) + (i-1)(i-2) + (n-1)(n-2) + (i+n-2)c}, \end{aligned} \tag{4.31}$$

$$\begin{aligned} S_5: \quad & \frac{q^{5+}(i)}{p^{5+}(i)} = \frac{q^{4-}(i)}{p^{4-}(i)} \\ & = \left(\frac{c^2 + (n-1)(i-1) + (i+n-2)c}{(n-1)(i-1) + (i-1)(i-2) + (n-1)(n-2) + (i+n-2)c} \right)^{-1}, \end{aligned} \tag{4.32}$$

$$\begin{aligned} S_6: \quad & \frac{q^{6+}(i)}{p^{6+}(i)} = \frac{q^{3-}(i)}{p^{3-}(i)} \\ & = \left(\frac{(n-1)(i-1) + (i+n-2)c}{c^2 + (n-1)(i-1) + (i-1)(i-2) + (n-1)(n-2) + (i+n-2)c} \right)^{-1}, \end{aligned} \tag{4.33}$$

$$S_7: \quad \frac{q^{7+}(i)}{p^{7+}(i)} = \frac{q^{2-}(i)}{p^{2-}(i)} = \left(\frac{c(N-2)}{(N-c-2)(N-3)} \right)^{-1}, \tag{4.34}$$

$$S_8: \quad \frac{q^{8+}(i)}{p^{8+}(i)} = \frac{q^{1-}(i)}{p^{1-}(i)} = \left(\frac{c(N-c-2)}{N^2 - N(c+5) + 3c + c^2 + 6} \right)^{-1}. \tag{4.35}$$

Unlike before, none of these are zero or one. We can, however, utilize another technique to determine the behavior. We will study the system's dynamics as a birth and death process for the population with trait value x, so that the probability of fixation ϕ can be calculated by means of a process that tracks the population of mutants (i) in the population of non-cheaters ($N - c$). Then, $i \in \{0, \ldots, N - c\}$, and the number of individuals with trait value y is given by $N - c - i$. States $i = 0$ (trait x is extinct) and $i = N - c$ (trait x has taken over) are absorbing. Given $p(i)$ and $q(i)$, we can determine the probability of fixation (i.e.

absorption in state $i = N - c$) by using the formula [25]

$$\frac{\sum_{k=1}^{N-c-1} \prod_{i=1}^{k} \frac{q(i)}{p(i)}}{1 + \sum_{k=1}^{N-c-1} \prod_{i=1}^{k} \frac{q(i)}{p(i)}},$$

which describes the probability of absorption into state zero from initial state 1. We are interested in the probability of being absorbed into state $N-c$, and since the only absorbing states are zero and $N-c$, a simple calculation gives us that the probability of absorption into state $N-c$ is

$$\left(1 + \sum_{k=1}^{N-c-1} \prod_{i=1}^{k} \frac{q(i)}{p(i)}\right)^{-1}. \tag{4.36}$$

For larger values of N, we can also use the *Mathematica* command DiscreteMarkovProc-cessto determine the probability of fixation, given a transition matrix. By performing the calculation in equation (4.36), we can exhaustively calculate the probability of fixation for all of the ratios in the equation set (4.35). In particular we note that S_1, S_2, S_7, S_8 do not depend on the number of mutants present, and so determining the probability of fixation is very easy.

In general, we will denote these probabilities of fixation as a_k, where the subscript refers to the case number from (4.27). These values are independent of the functional form of the payoff function or its parameters, but are dependent on the population sizes, N and c. The form of the payoff function instead determines the domains of the trait value in which these constants apply. To reflect the fact that different a_k's determine the evolution of the trait value as it changes, the function ϕ will be the sum of the a_k constants, multiplied by the appropriate Heaviside products, to produce a step function. For example, a_3 will be multiplied by

$$H\left(\mathcal{F}0(x-y)\right) H\left(\mathcal{F}(x-y)\right) H\left(-\left(\mathcal{F}+\mathcal{G}\right)\left(\mathcal{F}-\mathcal{G}\right)\right).$$

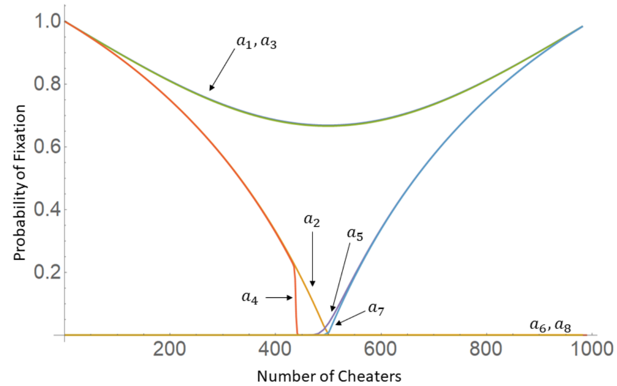


Figure 4.5: Probability of fixation for the different combinations of Heaviside functions as the number of cheaters changes in a population of $N = 1000$. Note the sharp decay of a_4.

This multiplier evaluates to one when both $\mathcal{F}0(x-y)$, and $\mathcal{F}(x-y)$ are positive and exactly one of $\mathcal{F}+\mathcal{G}$ and $\mathcal{F}-\mathcal{G}$ is positive; otherwise, it evaluates to zero. In this way, a_3 will describe the evolution of the trait value only when the trait value falls within case 3.

Figure 4.5 shows the behavior of a_k for each of the possible combinations, as a function of the number of cheaters, for a fixed value $N = 1000$. For the case when $N = 1000$, we find that only 6 of them result in a positive chance of fixation. In particular, $a_6 \equiv 0$ and $a_8 \equiv 0$. S_8 and S_1 have a complementary relationship, in which $q(i)/p(i)$ are reciprocals in the two cases. Since S_8 does not depend on the number of cheaters, a_8 does not either, so having $a_8 \equiv 0$ implies that $a_1 \neq 0$. S_6 and S_3 have a similar relationship, but $q(i)/p(i)$ in that case does depend on the number of mutants, so we can not draw the same conclusion. We can further observe from figure 4.5 and equations for the different cases that only one of a_2 and a_7 is positive, and similarly only one of a_4 and a_5 is positive (both could be zero).

These complementary behaviors are important as we move forward to write an expression for ϕ. As stated before, we can write $\phi(x, y)$ as a sum of coefficients a_k with appropriate Heaviside multipliers, and we can further simplify them by separating out $\Delta x = x - y$ from the Heaviside functions. We can concisely write

$$\frac{d\langle x\rangle}{dt} = \frac{\mu N\langle|\Delta x|\rangle}{2N}\int \mathcal{K}(y)P(y,t)dy, \tag{4.37}$$

where $\mathcal{K}$ depends on the signs of the partial derivatives and their sums and is defined below. For ease of notation, we will separate $\mathcal{K}(y)$ in to two equations. When $\mathcal{F}0 > 0$:

$$\mathcal{K}(y) = \begin{cases} a_1 & \text{if both } \mathcal{F}(y) \pm \mathcal{G}(y) > 0 \\ -a_2 & \text{if both } \mathcal{F}(y) \pm \mathcal{G}(y) < 0 \\ a_3 & \text{if } \mathcal{F}(y) > 0 \text{ and either } \mathcal{F}(y) + \mathcal{G}(y) > 0 \text{ or } \mathcal{F}(y) - \mathcal{G}(y) > 0 \\ -a_4 & \text{if } \mathcal{F}(y) < 0 \text{ and either } \mathcal{F}(y) + \mathcal{G}(y) > 0 \text{ or } \mathcal{F}(y) - \mathcal{G}(y) > 0 \\ a_5 & \text{if } \mathcal{F}(y) < 0 \text{ and either } \mathcal{F}(y) + \mathcal{G}(y) > 0 \text{ or } \mathcal{F}(y) - \mathcal{G}(y) > 0 \\ a_7 & \text{if both } \mathcal{F}(y) \pm \mathcal{G}(y) < 0 \\ 0 & \text{otherwise,} \end{cases} \tag{4.38}$$

and when $\mathcal{F}0 < 0$:

$$\mathcal{K}(y) = \begin{cases} -a_1 & \text{if both } \mathcal{F}(y) \pm \mathcal{G}(y) < 0 \\ a_2 & \text{if both } \mathcal{F}(y) \pm \mathcal{G}(y) > 0 \\ -a_3 & \text{if } \mathcal{F}(y) < 0 \text{ and either } \mathcal{F}(y) + \mathcal{G}(y) > 0 \text{ or } \mathcal{F}(y) - \mathcal{G}(y) > 0 \\ a_4 & \text{if } \mathcal{F}(y) > 0 \text{ and either } \mathcal{F}(y) + \mathcal{G}(y) > 0 \text{ or } \mathcal{F}(y) - \mathcal{G}(y) > 0 \\ -a_5 & \text{if } \mathcal{F}(y) > 0 \text{ and either } \mathcal{F}(y) + \mathcal{G}(y) > 0 \text{ or } \mathcal{F}(y) - \mathcal{G}(y) > 0 \\ -a_7 & \text{if both } \mathcal{F}(y) \pm \mathcal{G}(y) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.39}$$

The conditions in the right of the above expressions are exhaustive and cover all the possible combinations that are valid (recall that some are contradictory and were left out). Although these conditions are not mutually exclusive, this does not present a problem. In particular, note that the conditions for $-a_2$ and a_7 are the same, and the conditions for $-a_4$ and a_5 are the same. However, as discussed before, $-a_2$ and a_7 cannot both be non-zero for the same number of cheaters, and likewise for $-a_4$ and a_5, so we will not have to worry about this overlap.

As the final step, we will replace the mean value of a function by the function of the mean in equation (4.37) to arrive at the following equation for the mean trait:

$$\frac{d\langle x\rangle}{dt} = \frac{\mu N\langle|\Delta x|\rangle}{2N}\mathcal{K}(\langle x\rangle), \tag{4.40}$$

The choice for coefficients $\{a_1, \ldots, a_8\}$, can be expressed in words:

- The trait value moves at speed a_1 when the mutant does better with the resident, mutants or cheaters than the resident with those same individuals.
- It moves at speed a_2 or a_7 (depending on how many cheaters there are) if mutants do not do better with cheaters than residents with cheaters, but the mutants still perform better with residents or mutants than residents.
- The trait moves at speed a_3 if the mutant performs better than the resident with cheaters and the resident, but the resident is able to outperform mutants with either mutants or residents.
- The trait moves at speed a_4 or a_5 (depending on the number of cheaters) if the mutants perform better with residents than residents, but the resident is able to outperform mutants with either mutants or residents and the resident performs better with cheaters than the mutant.

From these descriptions, we can see that the number of cheaters does impact the behavior.

More cheaters can bring a trait value down, depending on how the residents and mutants interact with the cheaters.

4.6.3 Single Interaction: Probabilistic Criterion

We must again determine $\phi(x, y)$, the probability of fixation by a mutant with trait x in a population of individuals with trait y to use in the mean trait evolution equation, equation (4.2) when using the probabilistic criterion. To find the probability of fixation, we need to determine when an individual with trait x will replace one with trait y, which, under the probabilistic criterion, is proportional to the difference between the two payoff functions. So, replacement of an individual with trait y by an individual with trait x occurs when $(F(x, w) - F(y, z))/\alpha > r$, where w and z are trait values drawn from the population at random, α is a constant that ensures the difference will always be less than 1, and r is a random number drawn from a uniform distribution between 0 and 1.

For this scenario with cheaters, below we list all the case where an individual with trait x competes with an individual with trait y and wins by having a larger payoff (assuming the trait x is very close to the trait y):

$$F(x,0) - F(y,0) \approx \frac{\partial F(z_1,0)}{\partial z_1}|_{z=y}(x-y) > 0 \tag{4.41a}$$

$$F(x,y) - F(y,y) \approx \frac{\partial F(z_1,y)}{\partial z_1}|_{z=y}(x-y) > 0 \tag{4.41b}$$

$$F(x,x) - F(y,x) \approx \frac{\partial F(z_1,y)}{\partial z_1}|_{z=y}(x-y) > 0 \tag{4.41c}$$

$$F(x,x) - F(y,y) \approx \left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y} + \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\right)(x-y) > 0 \tag{4.41d}$$

$$F(x,y) - F(y,x) \approx \left(\frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y} - \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\right)(x-y) > 0 \tag{4.41e}$$

$$F(x,0) - F(y,y) > 0 \tag{4.41f}$$

$$F(x,0) - F(y,x) > 0 \tag{4.41g}$$

$$F(x,x) - F(y,0) > 0 \tag{4.41h}$$

$$F(x,y) - F(y,0) > 0 \tag{4.41i}$$

The probability of replacement disproportional to the difference between the payoffs, and is nonzero when this difference is positive. Therefore, we can use the functions in (4.41), multiplied by Heaviside functions to ensure positivity, to arrive at the probability that the number of individuals with trait x goes from i to $i+1$, $p(i)$, and the probability that the number of individuals with trait x goes from i to $i-1$, $q(i)$:

$$p(i) = \frac{in}{\alpha(N-c)(N-c-1)(N-2)(N-3)} \cdot$$
$$\Bigg(\begin{bmatrix} c(c-1) & (i-1)(n-1) & (i-1)(n-1) & ((i-1)(i-2)+(n-1)(n-2)) \end{bmatrix} \cdot$$
$$\begin{bmatrix} H\left(\frac{\partial F(z_1,0)}{\partial z_1}\big|_{z=y}(x-y)\right)\frac{\partial F(z_1,0)}{\partial z_1}\big|_{z=y} \\ H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}+\frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right)(x-y)\right)\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}+\frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right) \\ H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}-\frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right)(x-y)\right)\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}-\frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right) \\ H\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}(x-y)\right)\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y} \end{bmatrix}(x-y) \tag{4.42}$$
$$+(i-1)c(F(x,0)-F(y,x))H(F(x,0)-F(y,x))$$
$$+(i-1)c(F(x,x)-F(y,0))H(F(x,x)-F(y,0))$$
$$+(n-1)c(F(x,0)-F(y,y)H(F(x,0)-F(y,y))$$
$$+(n-1)c(F(x,y)-F(y,0))H(F(x,y)-F(y,0))\Bigg),$$

$$q(i) = \frac{in}{\alpha N(N-1)(N-2)(N-3)} \cdot$$
$$\Bigg(\begin{bmatrix} c(c-1) & (i-1)(n-1) & (i-1)(n-1) & ((i-1)(i-2)+(n-1)(n-2)) \end{bmatrix} \cdot$$
$$\begin{bmatrix} H\left(\frac{\partial F(z_1,0)}{\partial z_1}\big|_{z=y}(y-x)\right)\frac{\partial F(z_1,0)}{\partial z_1}\big|_{z=y} \\ H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}+\frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right)(y-x)\right)\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}+\frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right) \\ H\left(\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}-\frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right)(y-x)\right)\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}-\frac{\partial F(y,z_2)}{\partial z_2}\big|_{z_2=y}\right) \\ H\left(\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y}(y-x)\right)\frac{\partial F(z_1,y)}{\partial z_1}\big|_{z_1=y} \end{bmatrix}(y-x) \tag{4.43}$$
$$+(i-1)c(F(y,0)-F(x,x))H(F(y,0)-F(x,x))$$
$$+(i-1)c(F(y,x)-F(x,0))H(F(y,x)-F(x,0))$$
$$+(n-1)c(F(y,0)-F(x,y))H(F(y,0)-F(x,y))$$
$$+(n-1)c(F(y,y)-F(x,0))H(F(y,y)-F(x,0))\Bigg).$$

These probabilities are dependent on the payoff function; therefore, as before, we will examine the ratio between $p(i)$ and $q(i)$ for all the possible cases. We will again use condensed notation

$$\begin{aligned}\mathcal{F}0(y) &= \frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=y} \approx \frac{\partial F(z_1,0)}{\partial z_1}\Big|_{z_1=x}\\ \mathcal{F}(y) &= \frac{\partial F(z_1,x)}{\partial z_1}\Big|_{z_1=y} \approx \frac{\partial F(z_1,y)}{\partial z_1}\Big|_{z_1=y}\\ \mathcal{G}(y) &= \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=x} \approx \frac{\partial F(y,z_2)}{\partial z_2}\Big|_{z_2=y}\end{aligned}$$

There are four combinations of these functions present in the functions which determine the probability of fixation: $\mathcal{F}0$, $\mathcal{F}$, $\mathcal{F}+\mathcal{G}$ and $\mathcal{F}-\mathcal{G}$. As before, if $\mathcal{F}+\mathcal{G}$ is positive (negative) and $\mathcal{F}-\mathcal{G}$ is positive (negative), then $\mathcal{F}$ is also positive (negative). Unlike before, we can not group the cases when exactly one of $\mathcal{F}+\mathcal{G}$ and $\mathcal{F}-\mathcal{G}$ is positive, since now the magnitude of these values is used in the determination of the probability of fixation.

Given a particular form of the payoff function F, we could determine the probability of fixation using equation (4.36). Without this information on F, we cannot simplify further than

$$\frac{d\langle x\rangle}{dt} = \mu \int\int M(x-y)\phi(x,y)P(y,t)(x-y)dxdy. \tag{4.44}$$

If we are given a function F, $\phi(x,y)$ can be explicitly calculated along a grid of x and y using equations (4.36), (4.42) and (4.43). This grid can then be interpolated to create a surface in Mathematica, which is used in the calculation of the double integral in equation (4.44) to determine the dynamics of the mean trait value $\langle x\rangle$. To find the steady state, this function $\phi(x,y)$ is integrated. Figure 4.6 shows the surface $\phi(x,y)$ for the case where the payoff function is given by equation (4.45). Recall that in the single interaction case when using the deterministic criterion, if a trait value was invadable from above, that is, if $\phi(y+\epsilon,y) > 0$, then it was not invadable from below, and so we also had $\phi(y-\epsilon,y) = 0$.

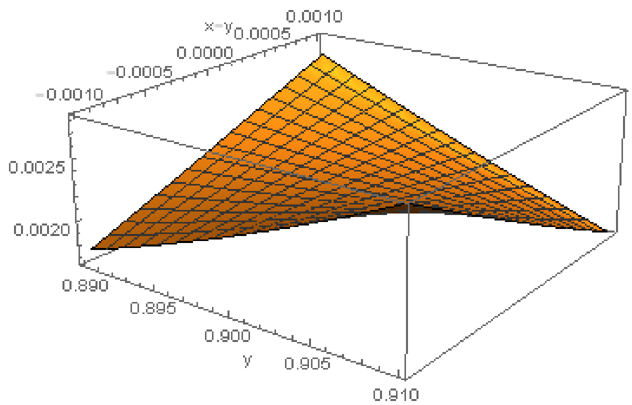


Figure 4.6: This surface is the graph of the function $\phi(x, y)$, where the payoff function is given by equation (4.45) and the parameter values are $\delta = .1$, $\gamma = .5$, $h = .5$, $g = .2$, and $m = 6$. For this calculation, $N = 1000$ and $c = 550$, and the steady state of equation (4.44) is calculated to be .899. The resident trait value is y while the mutant x is centered at the trait value y. This graph shows that it is possible for the probability of fixation to be positive for mutants with trait values both larger and smaller than the resident trait value.

As apparent from figure 4.6, this is not the case here. For the parameter values used in this simulation, the steady state of equation (4.44) corresponds to .899, and the probability of fixation by trait values both larger and smaller than the resident trait is positive. Since the probability of fixation by a mutant with either a larger and smaller trait value is positive at the steady state, upon reaching the steady state, the simulation is still susceptible to invasion by mutants both above and below the steady state. This will cause the system to fluctuate more about the steady state than the other models, see also comparisons in the next section (figure 4.7).

4.6.4 Comparison of the Models

To summarize our findings for this more complicated system that contains a population of cheaters, we note that we have considered four types of models:

- Multiple interactions, deterministic criterion. The ODE for the evolution of the mean trait value is given by equation (4.22).

- Multiple interactions, probabilistic criterion. The ODE is the same, equation (4.22).

- Single interaction, deterministic criterion. The equation is given by 4.37. This ODE contains different cases, which are listed in equations (4.38) and (4.39), and requires the calculation of fixation probabilities, $a_1, \ldots, a_7$, for Markov chains whose transition matrix information is summarized in equations (4.28-4.35). To implement this ODE, we wrote a function in MATLAB which determined the values of the various derivatives, and returned the appropriate a_k value. This was then used with a numerical solver to calculate the behavior of the ODE.

- Single interaction, probabilistic criterion. In this case, the method of determining the solution to the ODE is not dependent on a simplified form of equation (4.2). To calculate the dynamics then, $\phi(x, y)$ was explicitly calculated along a grid of x and y using equations (4.36), (4.42) and (4.43). This grid was then interpolated to create a surface in Mathematica, which was then used in the calculation of the double integral in equation (4.44) to determine the dynamics of the mean trait value $\langle x \rangle$.

For the comparisons of the different models, we will consider the particular payoff function:

$$F(x, y) = S_h(x + y)^2 - (2\delta x + \gamma S_g(x)^2), \tag{4.45}$$

where S_h and S_g are both Hill functions of the form:

$$S_h(x) = \frac{x^m}{x^m + h^m}.$$

Figure 4.7 demonstrates an example of the behavior of the four models where initially the population of non-cheaters is at trait .5, and the number of cheaters is 550 in a total population of 1000. We can see that not only the dynamics of convergence, but also the steady state depend on the model. We can see that, as predicted, the two models where the payoff is determined as the mean over multiple interactions, converge to the same equilibrium (the top two panels in figure 4.7). We further observe that the system with a single interaction and deterministic payoff (bottom left) converges to a different equilibrium. Since fluctuations are very small in this case, we can clearly see the difference in the behavior between this model and the models with multiple interactions. Finally, the last image (bottom right) corresponds to the system with a single interaction, but the probabilistic criterion. The value of the equilibrium (both predicted and numerically observed) is very similar to that for the deterministic criterion.

Therefore, we conclude that both the number of interactions that are included for the payoff calculation, and deterministic vs probabilistic replacement criterion, matter significantly for the mean trait dynamics. For the deterministic criterion, there is a difference between single and multiple interactions (compare the panels on the left in figure 4.7), and for a single interaction, there is a difference between deterministic and probabilistic criteria (compare the bottom panels in figure 4.7).

When we consider the case of the probabilistic criterion and compare single and multiple interactions (compare the panels on the right of figure 4.7) there is a very small difference in the ODE prediction for the location of the equilibrium. We however could not determine that the difference between the simulations in the two cases is significant.

Single interaction, deterministic vs probabilistic criterion. The differences in behavior of the different models arise due to the form of the probability of fixation function. Consider the two criteria for determining whether replacement happens. The probabilistic criterion uses the magnitude of the difference rather than just the sign. Because of this, a trait below or above the steady state has a nonzero probability to invade, as shown in figure 4.6. For the deterministic criterion, the probability of fixation only ever depends on

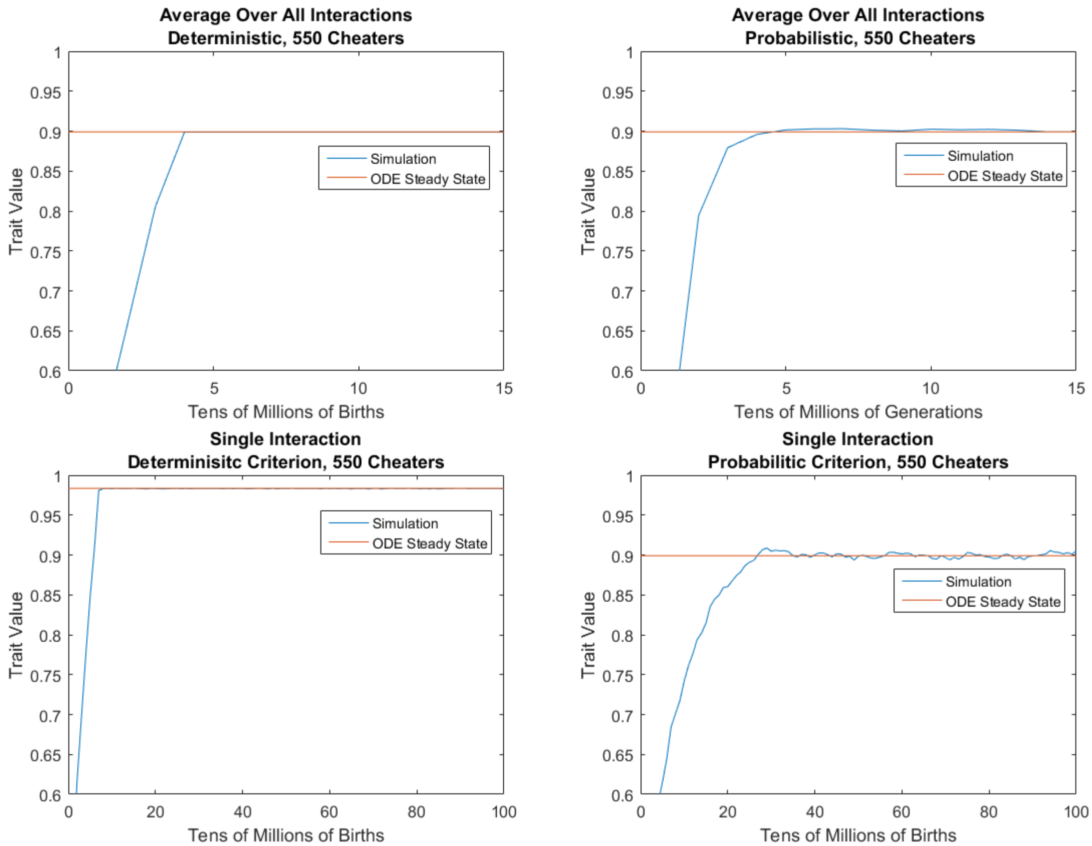


Figure 4.7: Comparison of the evolution of the traits when there are 550 cheaters, the payoff function is equation (4.45) with parameters $\delta = .1$, $\gamma = .5$, $h = .5$, $g = .2$, $m = 6$ and the population of non-cheaters start with a trait value of .5. The deterministic predictions of the equilibrium points are shown by horizontal lines. Top Left: The average of payoffs over all interactions was used with the deterministic criterion. Top Right: The average of payoffs over all interactions was used with the probabilistic criterion. In both these cases, The simulation is expected to converge to .8990. Bottom Left: Single interaction was used, with the deterministic criterion. The simulation is expected to converge to .9835. Bottom Right: Single interaction was used with the probabilistic criterion, which is expected to converge to .89898.

either the probability of invasion by smaller trait values or larger trait values, not both, as discussed in the derivation of equations (4.38) and (4.39). Both equations (4.38) and (4.39) are step functions. The change of value occurs when the dynamics of the trait value results in a change of case, see equations (4.27). This means that the equilibrium of the single interaction, deterministic criterion model must occur where there is a change in sign of $\mathcal{F}$, $\mathcal{F}+\mathcal{G}$, $\mathcal{F}-\mathcal{G}$, or $\mathcal{F}$. The single interaction, probabilistic criterion does not have such a restriction. It changes continuously (rather than as a piecewise constant) by allowing invasion by mutants with values higher and lower than the resident to invade, based on the magnitude of difference between the payoff functions (rather than only the sign).

Deterministic criterion, multiple vs single interactions. The reason for the difference is the behavior of the models with multiple and single interactions, for the deterministic criterion, is due to the different functions $p(i)$ and $q(i)$ that are generated. When many interactions are used, only one of $p(i)$ or $q(i)$ is non-zero. This is because $p(i)$ and $q(i)$ are Heaviside functions of the same argument, but with the opposite sign. When only one interaction is used, $p(i)$ and $q(i)$ can both be non-zero, since they are sums of Heaviside functions, rather than Heaviside functions of sums.

Probabilistic criterion, multiple vs single interactions. There are examples in which the ODEs for the mean trait evolution predict that the probabilistic method with a single interaction will converge to a significantly different value when compared to the simulations with many interactions. These examples however tend to occur when the probability of fixation from points above and below the current trait value are both likely. As shown in figure 4.6, the chance of mutation by larger and smaller trait values are both positive. The location of the steady state is where the probability of fixation of a larger trait value is equal to the probability of fixation by a smaller trait value. If an increase in the trait value happens, the probability of fixation by a lower trait value is larger, and vice versa. If the rate of fixation from both larger and smaller trait values is high at the steady state, the rate of convergence will be very slow, and the population will also appear rather unstable. Therefore, a numerical demonstration of statistical significance has so far been evasive.

4.7 Discussion and Conclusions

Formulating evolutionary models with selection and mutations is an important tool in mathematical biology. It not only provides a way to mathematically formulate and simulate behavior of particular biological systems, it also allows one to consider evolutionary patterns and "laws", to formulate evolutionary theories, and find explanations of various natural phenomena, both in natural evolution and other evolutionary applications.

It is therefore important to understand how general the conclusions drawn from a particular evolutionary model are. The particularities of models, the microscopic laws used in simulations, could be quite quite arbitrary, but are rarely discussed and are hardly mentioned in summaries of the results that were derived by using those models. It is usually assumed implicitly that those microscopic rules, if changed slightly and within reason, will lead to the same behavior. In this work we show, by using a particular evolutionary system, that these rules may or may not matter, and at least the variation of the rules must be investigated to understand just how general the results are.

We have chosen to focus on the evolutionary dynamics of a population of cooperators, whose fitness is determined by their interactions with other individuals. The idea is that cooperative interactions (or sharing some gene products) result in certain payoffs, which in turn translates into an increased propensity for reproduction. This verbal model can be translated into a simulation in a variety of ways. Here we concentrated on two binary choices: (i) In order to calculate the total payoff that translates into reproduction probability, individuals could interact with the whole population (that is, participate in numerous pairwise interactions, with the total payoff determining the fitness), or they could interact with a restricted set of randomly chosen individuals; to make the difference (if any) between these scenarios larger, we considered the extreme case of just one interaction. (ii) Once the fitness of a given individual is calculated, how do we decide if this individual will leave offspring? One possibility is to compare the total payoff of the focal individual with that of a competitor, and the individual who wins this competition will leave offspring (this is what we called the deterministic criterion). Alternatively, one could assume that the focus individual will be replaced by the offspring of a competitor only probabilistically, with the probability proportional to the fitness difference (the probabilistic criterion).

Superficially, such details should not matter. Whether an individual gets a chance to interact with only one other individual, or if they are paired off with every other individual before the fitness is determined, should somehow average out once we consider a large population of individuals. Also, since both competition rules look reasonable, one does not expect this to make a difference. To check whether this is true, we derived the ODEs governing the evolution of the mean trait. The detailed assumptions of the model all contribute to the derivation. As expected, in the simplest case of a population of cooperators with one evolving trait, no difference was found among the 4 cases, and all were shown to be described by the same ODE.

The results, however, were quite different once we slightly increased the complexity of the system and included a subpopulation of "cheaters". These individuals were able to take advantage of the products shared by the cooperators, but did not produce any of their own. In this case, there were more scenarios that could potentially occur, and consequently, the microscopic rules began to matter. We have shown that although the model in which interactions with the whole population contributed to the payoff, the deterministic vs probabilistic criterion did not make a difference, the model in which only a single individual contributed to the payoff, the deterministic criterion resulted in a different value of the steady state. In other words, we can say that under the deterministic update criterion, the number of interactions matters, and under the assumption of a single interaction, the choice of either the probabilistic or deterministic criterion does make a significant difference.

Bibliography

[1] R. Axelrod, D. E. Axelrod, and K. J. Pienta. Evolution of cooperation among tumor cells. *Proceedings of the National Academy of Sciences*, 103(36):13474–13479, 2006.

[2] P. B. Beauregard, Y. Chai, H. Vlamakis, R. Losick, and R. Koltera. Bacillus subtilis biofilm induction by plant polysaccharides. *Proc Natl Acad Sci USA*, 110:E1621–E1630, 2013.

[3] Å. Brännström and U. Dieckmann. Evolutionary dynamics of altruism and cheating among social amoebas. *Proc. Biol. Sci.*, 272(1572):1609–1616, Aug 2005.

[4] Å. Brännström, J. Johansson, and N. von Festenberg. The hitchhikers guide to adaptive dynamics. *Games*, 4(3):304, 2013.

[5] M. A. Brockhurst, A. Buckling, and A. Gardner. Cooperation peaks at intermediate disturbance. *Current Biology*, 17:761–765, 2007.

[6] M. A. Brockhurst, M. E. Hochberg, T. Bell, and A. Buckling. Character displacement promotes cooperation in bacterial biofilms. *Current biology*, 16(20):2030–2034, 2006.

[7] J. L. Bronstein. The evolution of facilitation and mutualism. *Journal of Ecology*, 97(6):1160–1170, 2009.

[8] L. W. Buss. *The evolution of individuality*. Princeton University Press, 2014.

[9] R. Connor. Invested, extracted and byproduct benefits: a modified scheme for the evolution of cooperation. *Behavioural Processes*, 76(2):109–113, 2007.

[10] T. Day and K. A. Young. Competitive and facilitative evolutionary diversification. *BioScience*, 54(2):101–109, 2004.

[11] F. Dercole and S. Rinaldi. *Analysis of evolutionary processes: the adaptive dynamics approach and its applications*. Princeton University Press, 2008.

[12] U. Dieckmann and R. Law. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J Math Biol*, 34(5-6):579–612, 1996.

[13] M. Dobeli. *Adaptive Diversification.* Princeton University Press, New Jersey, USA, 2011.

[14] M. Doebeli. A model for the evolutionary dynamics of cross-feeding polymorphisms in microorganisms. *Population Ecology*, 44(2):59–70, 2002.

[15] M. Doebeli and U. Dieckmann. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *The american naturalist*, 156(S4):S77–S101, 2000.

[16] M. Doebeli, C. Hauert, and T. Killingback. The evolutionary origin of cooperators and defectors. *Science*, 306(5697):859–862, Oct 2004.

[17] S. Estrela and I. Gudelj. Evolution of cooperative cross-feeding could be less challenging than originally thought. *PloS one*, 5(11):e14121, 2010.

[18] S. Gavrilets. Rapid transition towards the division of labor via evolution of developmental plasticity. *PLoS Comput Biol*, 6(6):e1000805, 2010.

[19] S. A. Geritz, G. Mesze, J. A. Metz, et al. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary ecology*, 12(1):35–57, 1998.

[20] S. A. Geritz, J. A. Metz, and C. Rueffler. Mutual invadability near evolutionarily singular strategies for multivariate traits, with special reference to the strongly convergence stable case. *Journal of mathematical biology*, 72(4):1081–1099, 2016.

[21] M. Gyllenberg and K. Parvinen. Necessary and sufficient conditions for evolutionary suicide. *Bulletin of mathematical biology*, 63(5):981–993, 2001.

[22] M. Gyllenberg, K. Parvinen, and U. Dieckmann. Evolutionary suicide and evolution of dispersal in structured metapopulations. *Journal of mathematical biology*, 45(2):79–105, 2002.

[23] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *cell*, 144(5):646–674, 2011.

[24] W. Harder. Enrichment and characterization of degrading organisms. In T. Leisinger, R. Htter, A. Cook, and J. Nesch, editors, *Microbial degradation of xenobiotics and recalcitrant compounds*, FEMS Symposium 12, pages 77–96. Academic Press, London, 1981.

[25] S. Karlin. *A first course in stochastic processes*. Academic press, 2014.

[26] W. Kim, S. B. Levy, and K. R. Foster. Rapid radiation in bacteria leads to a division of labour. *Nature communications*, 7, 2016.

[27] N. L. Komarova, L. Shahriyari, and D. Wodarz. Complex role of space in the crossing of fitness valleys by asexual populations. *Journal of The Royal Society Interface*, 11(95):20140014, 2014.

[28] N. L. Komarova, E. Urwin, and D. Wodarz. Accelerated crossing of fitness valleys through division of labor and cheating in asexual populations. *Scientific reports*, 2:917, 2012.

[29] A. Kroon and C. Van Ginkel. Complete mineralization of dodecyldimethylamine using a two-membered bacterial culture. *Environmental microbiology*, 3(2):131–136, 2001.

[30] O. Leimar. Evolutionary change and darwinian demons. *Selection*, 2(1-2):65–72, 2002.

[31] O. Leimar. The evolution of pheontypic polymorphism: Randomized strategies versus evolutionary branching. *Am Nat*, 165:669–681, 2005.

[32] O. Leimar. Multidimensional convergence stability. *Evolutionary Ecology Research*, 11(2):191–208, 2009.

[33] D. Lopez, H. Vlamakis, and R. Kolter. Generation of multiple cell types in bacillus subtilis. *FEMS Microbiology Reviews*, 33(1):152–163, 2009.

[34] B. J. McGill and J. S. Brown. Evolutionary game theory and adaptive dynamics of continuous traits. *Annual Review of Ecology, Evolution, and Systematics*, pages 403–435, 2007.

[35] D. O. Meltzer and J. W. Chung. Coordination, switching costs and the division of labor in general medicine: An economic explanation for the emergence of hospitalists in the united states. Working Paper 16040, National Bureau of Economic Research, May 2010.

[36] J. A. Metz, R. M. Nisbet, and S. A. Geritz. How should we define 'fitness' for general ecological scenarios? *Trends Ecol. Evol. (Amst.)*, 7(6):198–202, Jun 1992.

[37] J. J. Morris, R. E. Lenski, and E. R. Zinser. The black queen hypothesis: evolution of dependencies through adaptive gene loss. *MBio*, 3(2):e00036–12, 2012.

[38] M. A. Nowak. Five rules for the evolution of cooperation. *science*, 314(5805):1560–1563, 2006.

[39] N. M. Oliveira, R. Niehus, and K. R. Foster. Evolutionary limits to cooperation in microbial communities. *Proceedings of the National Academy of Sciences*, 111(50):17941–17946, 2014.

[40] T. Pfeiffer and S. Bonhoeffer. Evolution of cross-feeding in microbial populations. *The American naturalist*, 163(6):E126–E135, 2004.

[41] P. B. Rainey. Bacteria evolve and function within communities: observations from experimental pseudomonas populations. In M. J. McFail-Ngai, B. Henderson, and N. Ruby, editors, *The Influence of Cooperative Bacteria on Animal Host Biology*, pages 83–102. Cambridge University Press, 2005.

[42] P. B. Rainey and K. Rainey. Evolution of cooperation and conflict in experimental bacterial populations. *Nature*, 425(6953):72–74, 2003.

[43] N. Ribeck and R. E. Lenski. Modeling and quantifying frequency-dependent fitness in microbial populations with cross-feeding interactions. *Evolution*, 69(5):1313–1320, 2015.

[44] V. Rossetti, B. E. Schirrmeister, M. V. Bernasconi, and H. C. Bagheri. The evolutionary path to terminal differentiation and division of labor in cyanobacteria. *Journal of Theoretical Biology*, 262(1):23–34, 2010.

[45] C. Rueffler, J. Hermisson, and G. P. Wagner. Evolution of functional specialization and division of labor. *Proceedings of the National Academy of Sciences*, 109(6):E326–E335, 2012.

[46] W. Shou, S. Ram, and J. M. Vilar. Synthetic cooperation in engineered yeast populations. *Proceedings of the National Academy of Sciences*, 104(6):1877–1882, 2007.

[47] A. Smith. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Oxford: Claredon Press, reprinted 1976 edition, 1776.

[48] R. Sugden. *The Economics of Rights, Co-operation and Welfare.* Blackwell, Oxford, 1986.

[49] P. E. Turner, V. Souza, and R. E. Lenski. Tests of ecological mechanisms promoting the stable coexistence of two bacterial genotypes. *Ecology*, 77(7):2119–2129, 1996.

[50] Z. Vásárhelyi, G. Meszéna, and I. Scheuring. Evolution of heritable behavioural differences in a model of social division of labour. *PeerJ*, 3:e977, 2015.

[51] L. Wahl. Evolving the division of labour: generalists, specialists and task allocation. *Journal of Theoretical Biology*, 219(3):371–388, 2002.

[52] J. B. Xavier and K. R. Foster. Cooperation and conflict in microbial biofilms. *Proceedings of the National Academy of Sciences*, 104(3):876–881, 2007.

Appendix A

Supplementary Material for Chapter 2

A.1 Mutations in a Two-Population Quasispecies System

We would like to examine how mutations in a quasispecies system affect the steady states of the system. To this end, we will examine a two-population system where mutation is possible between the two populations.

First, we will examine the behavior of the two cell populations without mutation. In system A.1, the growth rates of populations x_1 and x_2 are r_1 and r_2, respectively.

$$\dot{x_1} = r_1 x_1 - (r_1 x_1 + r_2 x_2) x_1 \tag{A.1a}$$

$$\dot{x_2} = r_2 x_2 - (r_1 x_1 + r_2 x_2) x_2 \tag{A.1b}$$

The steady states are $(x_1, x_2) = (1, 0)$ when $r_1 > r_2$ and $(x_1, x_2) = (0, 1)$ when $r_1 < r_2$. If $r_1 = r_2$ then the populations remain at the initial conditions.

Now, we will examine the same system allowing for mutations between the two populations. We will denote the mutation rate from the first population to the second by u, while the rate from the second to the first is v. This interaction is described by system A.2.

$$\dot{x_1} = r_1x_1(1-u) + r_2x_2v - (r_1x_1 + r_2x_2)x_1 \tag{A.2a}$$

$$\dot{x_2} = r_1x_1u + r_2x_2(1-v) - (r_1x_1 + r_2x_2)x_2 \tag{A.2b}$$

Treating u and v as variables near zero, a Taylor expansion is applied to the steady states. If $r_1 > r_2$ then the following steady state is stable.

$$x_1 \to 1 + O(v^2) + \left(\frac{-r_1}{r_1 - r_2} + \frac{r_1r_2v}{(r_1-r_2)^2} + O(v^2)\right)u + O(u^2)$$
$$x_2 \to O(v^2) + \left(\frac{r_1}{r_1 - r_2} - \frac{r_1r_2v}{(r_1-r_2)^2} + O(v^2)\right)u + O(u^2).$$

There is a second solution, which is stable for $r_2 > r_1$:

$$x_1 \to \frac{-r_2v}{r_1 - r_2} + O(v^2) + \left(\frac{-r_1r_2v}{(r_1-r_2)^2} + O(v^2)\right)u + O(u^2)$$
$$x_2 \to 1 + \frac{r_2v}{r_1 - r_2} + O(v^2) + \left(\frac{r_1r_2v}{(r_1-r_2)^2} + O(v^2)\right)u + O(u^2)$$

We can recognize both of these solutions as small perturbations of the steady states of system A.1 (no mutations). Thus, if the mutation rates can be assumed to be small, then knowing the steady states in the non-mutated system can give insight to the mutated steady states.

A.2 Eigenvalue Analysis for Chapter 2

A.2.1 Asymmetric Cooperation: Linear Cooperation Function

Consider the system below:

$$\dot{x_{ab}} = r_{ab}x_{ab} - \phi x_{ab} \tag{A.3a}$$

$$\dot{x_{aB}} = r_{aB}x_{aB} - \phi x_{aB} \tag{A.3b}$$

$$\dot{x_{Ab}} = [r_{Ab} + \alpha(x_{aB} + x_{AB})]x_{Ab} - \phi x_{Ab} \tag{A.3c}$$

$$\dot{x_{AB}} = r_{AB}x_{AB} - \phi x_{AB}. \tag{A.3d}$$

The steady states and eigenvalues of the Jacobian of the system are listed below.

1. The eigenvalues when $(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (1, 0, 0, 0)$ are

$$\lambda_1 = -r_{ab} \qquad \lambda_2 = -r_{ab} + r_{aB}$$

$$\lambda_3 = -r_{ab} + r_{Ab} \qquad \lambda_4 = -r_{ab} + r_{AB}.$$

   By assuming that growth rates are positive, this state will be stable when $r_{ab} > r_{aB}$, $r_{ab} > r_{Ab}$, $r_{ab} > r_{AB}$.

2. The eigenvalues when $(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 0, 0, 1)$ are

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = r_{aB} - r_{AB}$$

$$\lambda_3 = \alpha + r_{Ab} - r_{AB} \qquad \lambda_4 = -r_{AB}.$$

   By assuming that the growth rates are positive, this state will be stable when $r_{AB} > r_{ab}$, $r_{AB} > r_{aB}$, $r_{AB} > \alpha + r_{Ab}$.

3. The eigenvalues when $(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 1, 0, 0)$ are

$$\lambda_1 = r_{ab} - r_{aB} \qquad \lambda_2 = -r_{aB}$$

$$\lambda_3 = \alpha - r_{aB} + r_{Ab} \qquad \lambda_4 = -r_{aB} + r_{AB}.$$

By assuming that the growth rates are positive, this state will be stable when $r_{aB} > r_{ab}$, $r_{aB} > \alpha + r_{Ab}$, $r_{aB} > r_{AB}$.

4. The eigenvalues when $(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 0, 1, 0)$ are

$$\lambda_1 = r_{ab} - r_{Ab} \qquad \lambda_2 = r_{aB} - r_{Ab}$$

$$\lambda_3 = -r_{Ab} \qquad \lambda_4 = -r_{Ab} + r_{AB}.$$

By assuming growth rates are positive, this state will be stable when $r_{Ab} > r_{ab}$, $r_{Ab} > r_{aB}$, $r_{Ab} > r_{AB}$.

5. The eigenvalues when $(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, \frac{r_{aB}-r_{Ab}}{\alpha}, \frac{\alpha - r_{aB}+r_{Ab}}{\alpha}, 0)$ are

$$\lambda_1 = r_{ab} - r_{aB} \qquad \lambda_2 = -r_{aB}$$

$$\lambda_3 = -\frac{(r_{aB} - r_{Ab})(\alpha - r_{aB} + r_{Ab})}{\alpha} \qquad \lambda_4 = -r_{aB} + r_{AB}.$$

By assuming growth rates are positive, $\lambda_2 < 0$. Note also that by assuming that the populations and the parameter α are positive, we have that $r_{aB} > r_{Ab}$ and $\alpha + r_{Ab} > r_{aB}$, so $\lambda_3 < 0$. Thus, by assuming positive populations, positive parameters and positive growth rates, this state will be stable when $r_{aB} > r_{ab}$, $r_{aB} > r_{AB}$.

6. The eigenvalues when $(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 0, \frac{r_{AB}-r_{Ab}}{\alpha}, \frac{\alpha+r_{Ab}-r_{AB}}{\alpha})$ are

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = r_{aB} - r_{AB}$$

$$\lambda_3 = \frac{(r_{Ab} - r_{AB})(\alpha + r_{Ab} - r_{AB})}{\alpha} \qquad \lambda_4 = -r_{AB}.$$

By assuming the growth rates are positive, $\lambda_4 = -r_{AB} < 0$. Note that by also assuming the populations and α are positive, we have that $r_{Ab} - r_{AB} < 0$, and $\alpha + r_{Ab} - r_{AB} > 0$, so $\lambda_3 < 0$. So, when requiring the populations, parameter α, and growth rates to be positive, this state will be stable if $r_{AB} > r_{ab}$, $r_{AB} > r_{aB}$.

A.2.2 Linear Cooperation Function, Three Populations

Consider the system below:

$$\dot{x_{ab}} = r_{ab}x_{ab} - \phi x_{ab} \tag{A.4a}$$

$$\dot{x_{Ab}} = r_{Ab}x_{Ab} + \alpha x_{aB}x_{Ab} - \phi x_{Ab} \tag{A.4b}$$

$$\dot{x_{aB}} = r_{aB}x_{aB} + \alpha x_{Ab}x_{aB} - \phi x_{aB}. \tag{A.4c}$$

The steady states, eigenvalues of the Jacobian, and the stability of the system are discussed below.

1. The eigenvalues when $(x_{ab}, x_{Ab}, x_{aB}) = (1, 0, 0)$ are

$$\lambda_1 = r_{ab}$$

$$\lambda_2 = -r_{ab} + r_{Ab}$$

$$\lambda_3 = -r_{ab} + r_{aB}.$$

Assuming positive growth rates, this state will be stable if $r_{ab} > r_{Ab}$ and $r_{ab} > r_{aB}$.

2. The eigenvalues when $(x_{ab}, x_{Ab}, x_{aB}) = (0, 1, 0)$ are

$$\lambda_1 = r_{ab} - r_{Ab}$$
$$\lambda_2 = -r_{Ab}$$
$$\lambda_3 = \alpha - r_{Ab} + r_{aB}.$$

Assuming positive growth rates, this state will be stable if $r_{Ab} > r_{ab}$ and $r_{Ab} > \alpha + r_{aB}$.

3. The eigenvalues when $(x_{ab}, x_{Ab}, x_{aB}) = (0, 0, 1)$ are

$$\lambda_1 = r_{ab} - r_{aB}$$
$$\lambda_2 = \alpha + r_{Ab} - r_{aB}$$
$$\lambda_3 = -r_{aB}.$$

Assuming positive growth rates, this state will be stable if $r_{aB} > r_{ab}$ and $r_{aB} - r_{Ab} > \alpha$.

4. The eigenvalues when $(x_{ab}, x_{Ab}, x_{aB}) = (\frac{\alpha - 2r_{ab} + r_{Ab} + r_{aB}}{\alpha}, \frac{r_{ab} - r_{aB}}{\alpha}, \frac{r_{ab} - r_{Ab}}{\alpha})$ are

$$\begin{aligned}
\lambda_1 &= -r_{ab} \\
\lambda_2 &= -\frac{1}{\alpha^3}\Bigg(\alpha^2(r_{ab} - r_{Ab})(r_{ab} - r_{aB}) \\
&\qquad + \sqrt{\alpha^4(r_{ab} - r_{Ab})(\alpha - r_{ab} + r_{Ab})(r_{ab} - r_{aB})(\alpha - r_{ab} + r_{aB})}\Bigg) \\
&= -\frac{1}{\alpha^3}\left(\gamma + \sqrt{\gamma^2 + \delta}\right) \\
\lambda_3 &= -\frac{1}{\alpha^3}\Bigg(\alpha^2(r_{ab} - r_{Ab})(r_{ab} - r_{aB}) \\
&\qquad - \sqrt{\alpha^4(r_{ab} - r_{Ab})(\alpha - r_{ab} + r_{Ab})(r_{ab} - r_{aB})(\alpha - r_{ab} + r_{aB})}\Bigg) \\
&= -\frac{1}{\alpha^3}\left(\gamma - \sqrt{\gamma^2 + \delta}\right)
\end{aligned}$$

where

$$\gamma = \alpha^2(r_{ab} - r_{Ab})(r_{ab} - r_{aB})$$
$$\delta = \alpha^5(r_{ab} - r_{Ab})(r_{ab} - r_{aB})(\alpha - 2r_{ab} + r_{Ab} + r_{aB}).$$

If we assume positive growth rates, λ_1 is negative. λ_2 will have a negative real part if $(r_{ab} - r_{Ab})(r_{ab} - r_{aB})$ is positive, and λ_3 will have negative real part if $\delta < 0$. Starting with λ_3, if $\delta < 0$, there are two possibilities. Either $\sqrt{\gamma^2 + \delta}$ is less than γ, so that $\gamma - \sqrt{\gamma^2 + \delta}$ is positive, making λ_3 negative, or $\sqrt{\gamma^2 + \delta}$ is imaginary, which doesn't affect the stability. δ is negative if $(r_{ab} - r_{Ab})(r_{ab} - r_{aB})(\alpha - 2r_{ab} + r_{Ab} + r_{aB}) < 0$. If we assume that $\lambda_2 < 0$, then it is already the case that $(r_{ab} - r_{Ab})(r_{ab} - r_{aB}) > 0$, so for λ_3 we are left with the requirement that $(\alpha - 2r_{ab} + r_{Ab} + r_{aB}) < 0$. We are interested in positive populations, and by examining the equations to describe the populations, we see that each of $(r_{ab} - r_{Ab})$, $(r_{ab} - r_{aB})$ and $(\alpha - 2r_{ab} + r_{Ab} + r_{aB})$ must be positive. Thus, the positivity requirement on the population eliminates the possibility of stability.

5. The eigenvalues when $(x_{ab}, x_{Ab}, x_{aB}) = (0, \frac{\alpha + r_{Ab} - r_{aB}}{2\alpha}, \frac{\alpha - r_{Ab} + r_{aB}}{2\alpha})$ are

$$\lambda_1 = \frac{1}{2}(-\alpha - r_{Ab} - r_{aB})$$
$$\lambda_2 = \frac{1}{2}(-\alpha + 2r_{ab} - r_{Ab} - r_{aB})$$
$$\lambda_3 = \frac{-\alpha^2 + (r_{Ab} - r_{aB})^2}{2\alpha}.$$

This state will be stable if $2r_{ab} < \alpha + r_{Ab} + r_{aB}$ and $(r_{Ab} - r_{aB})^2 - \alpha^2 < 0$. Note that

$$\begin{aligned}(r_{Ab} - r_{aB})^2 - \alpha^2 &= (r_{Ab} - r_{aB} - \alpha)(r_{Ab} - r_{aB} + \alpha) \\ &= (-\alpha + r_{Ab} - r_{aB})(\alpha + r_{Ab} - r_{aB}) \\ &= -(\alpha - r_{Ab} + r_{aB})(\alpha + r_{Ab} - r_{aB}).\end{aligned}$$

If the populations are positive, then

$$-(\alpha - r_{Ab} + r_{aB})(\alpha + r_{Ab} - r_{aB}) < 0.$$

Thus, if the populations are positive and $2r_{ab} < \alpha + r_{Ab} + r_{aB}$ this state is stable.

A.2.3 Linear Cooperation Functions, Four Populations

Consider the system below,

$$\dot{x_{ab}} = r_{ab}x_{ab} - \phi x_{ab} \tag{A.5a}$$

$$\dot{x_{aB}} = [r_{aB} + \beta(x_{Ab} + x_{AB})]x_{aB} - \phi x_{aB} \tag{A.5b}$$

$$\dot{x_{Ab}} = [r_{Ab} + \alpha(x_{aB} + x_{AB})]x_{Ab} - \phi x_{Ab} \tag{A.5c}$$

$$\dot{x_{AB}} = r_{AB}x_{AB} - \phi x_{AB}. \tag{A.5d}$$

The steady states and eigenvalues of the Jacobian of the system are listed below.

1. The eigenvalues when

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (1, 0, 0, 0)$$

are

$$\lambda_1 = -r_{ab} \qquad \lambda_2 = -r_{ab} + r_{aB}$$

$$\lambda_3 = -r_{ab} + r_{Ab} \qquad \lambda_4 = -r_{ab} + r_{AB}$$

Assuming positive growth rates, this state will be stable if $r_{ab} > r_{aB}$, $r_{ab} > r_{Ab}$ and $r_{ab} > r_{AB}$.

2. The eigenvalues when

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 1, 0, 0)$$

are

$$\lambda_1 = r_{ab} - r_{aB} \qquad \lambda_2 = -r_{aB}$$

$$\lambda_3 = \alpha - r_{aB} + r_{Ab} \qquad \lambda_4 = -r_{aB} + r_{AB}.$$

Assuming positive growth rates, this state will be stable if $r_{aB} > r_{ab}$, $r_{aB} > \alpha + r_{Ab}$ and $r_{aB} > r_{AB}$.

3. The eigenvalues when

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 0, 1, 0)$$

are

$$\lambda_1 = r_{ab} - r_{Ab} \qquad \lambda_2 = \beta + r_{aB} - r_{Ab}$$

$$\lambda_3 = -r_{Ab} \qquad \lambda_4 = -r_{Ab} + r_{AB}.$$

Assuming positive growth rates, this state will be stable if $r_{Ab} > r_{ab}$, $r_{Ab} > \beta + r_{aB}$ and $r_{Ab} > r_{AB}$.

4. The eigenvalues when

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 0, 0, 1)$$

are

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = \beta + r_{aB} - r_{AB}$$

$$\lambda_3 = \alpha + r_{Ab} - r_{AB} \qquad \lambda_4 = -r_{AB}.$$

Assuming positive growth rates, this state will be stable if $r_{AB} > r_{ab}$, $r_{AB} > \beta + r_{aB}$ and $r_{AB} > \alpha + r_{Ab}$.

5. The eigenvalues when

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = \left(0, \frac{\beta + r_{aB} - r_{Ab}}{\alpha + \beta}, \frac{\alpha - r_{aB} + r_{Ab}}{\alpha + \beta}, 0\right)$$

are

$$\lambda_1 = -\frac{(\beta + r_{aB} - r_{Ab})(\alpha - r_{aB} + r_{Ab})}{\alpha + \beta} \qquad \lambda_2 = -\frac{\alpha(\beta + r_{aB}) + \beta r_{Ab}}{\alpha + \beta}$$

$$\lambda_3 = -\frac{\alpha(\beta - r_{ab} + r_{aB}) - \beta(r_{ab} - r_{Ab})}{\alpha + \beta} \qquad \lambda_4 = -\frac{\alpha(\beta + r_{aB}) + \beta r_{Ab}}{\alpha + \beta} + r_{AB}.$$

λ_1 is negative when the populations are positive since the numerator is the product of the numerators of the non-zero populations. λ_2 is negative when $\alpha(\beta + r_{ab}) + \beta r_{Ab} > 0$, which is always the case since the parameters are all assumed to be positive. λ_3 is negative when $\alpha(\beta - r_{ab} + r_{aB}) - \beta(r_{ab} - r_{Ab}) > 0$. The fourth eigenvalue, λ_4 is negative when $\frac{\alpha(\beta + r_{aB}) + \beta r_{Ab}}{\alpha + \beta} > r_{AB}$. Thus the two requirements we have for stability, other than positive populations and positive parameters, are then

$$\alpha(\beta - r_{ab} + r_{aB}) - \beta(r_{ab} - r_{Ab}) > 0$$
$$\frac{\alpha(\beta + r_{aB}) + \beta r_{Ab}}{\alpha + \beta} > r_{AB}.$$

Notice that if $\lambda_4 < 0$, we can say

$$\begin{aligned}
& \frac{\alpha(\beta + r_{aB}) + \beta r_{Ab}}{\alpha + \beta} > r_{AB} \\
\Rightarrow \quad & \frac{\alpha(\beta + r_{aB}) + \beta r_{Ab}}{\alpha + \beta} + \frac{-\alpha r_{ab} - \beta r_{ab}}{\alpha + \beta} > \frac{-\alpha r_{ab} - \beta r_{ab}}{\alpha + \beta} + r_{AB} \\
\Rightarrow \quad & \frac{\alpha(\beta - r_{ab} + r_{aB}) - \beta(r_{ab} - r_{Ab})}{\alpha + \beta} > r_{AB} - r_{ab}.
\end{aligned}$$

If $r_{AB} > r_{ab}$, then satisfying $\lambda_4 < 0$ satisfies $\lambda_3 < 0$.

6. The eigenvalues when

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = \left(0, \frac{\beta + r_{aB} - r_{AB}}{\beta}, 0, \frac{-r_{aB} + r_{AB}}{\beta}\right)$$

are

$$\begin{aligned}
\lambda_1 &= r_{ab} - r_{AB} & \lambda_2 &= -\frac{(-r_{aB} + r_{AB})(\beta + r_{aB} - r_{AB})}{\beta} \\
\lambda_3 &= \alpha + r_{Ab} - r_{AB} & \lambda_4 &= -r_{AB}.
\end{aligned}$$

Assuming positive growth rates and $\beta > 0$, this state is stable if $r_{AB} > r_{ab}$, $r_{AB} > \alpha + r_{Ab}$ and $(-r_{aB} + r_{AB})(\beta + r_{aB} - r_{AB}) > 0$. Note that the numerator of λ_2 is $-\beta x_{aB} x_{AB}$ for the steady state values of x_{aB} and x_{AB}. Thus, if the populations are positive, then λ_2 is negative. Thus, the requirements for stability given positive parameters and positive populations reduce to $r_{ab} < r_{AB}$ and $\alpha + r_{Ab} < r_{AB}$.

7. The eigenvalues when

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = \left(0, 0, \frac{\alpha + r_{Ab} - r_{AB}}{\alpha}, \frac{-r_{Ab} + r_{AB}}{\alpha}\right)$$

are

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = \beta + r_{aB} - r_{AB}$$
$$\lambda_3 = \frac{(r_{Ab} - r_{AB})(\alpha + r_{Ab} - r_{AB})}{\alpha} \qquad \lambda_4 = -r_{AB}.$$

Assuming positive growth rates and $\alpha > 0$, this state will be stable if $r_{AB} > r_{ab}$, $r_{AB} > \beta + r_{aB}$ and $(r_{Ab} - r_{AB})(\alpha + r_{Ab} - r_{AB}) < 0$. The third condition is satisfied by requiring the populations to be positive. This state is stable for $r_{ab} < r_{AB}$ and $\beta + r_{aB} < r_{AB}$.

8. The eigenvalues when

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = \left(0, \frac{\beta + r_{aB} - r_{AB}}{\beta}, \frac{\alpha + r_{Ab} - r_{AB}}{\alpha}, \frac{-\alpha(\beta + r_{aB} - r_{AB}) + \beta(-r_{Ab} + r_{AB})}{\alpha\beta}\right)$$

are

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = -r_{AB}$$
$$\lambda_3 = -\frac{1}{2\alpha\beta}\left(\gamma + \sqrt{\gamma^2 + \delta}\right) \qquad \lambda_4 = -\frac{1}{2\alpha\beta}\left(\gamma - \sqrt{\gamma^2 + \delta}\right).$$

where

$$\gamma = -\alpha\beta(r_{aB} + r_{Ab} - 2r_{AB}) - \alpha(r_{aB} - r_{AB})^2 - \beta(r_{Ab} - r_{AB})^2$$
$$\delta = 4\alpha\beta(\alpha(\beta + r_{aB} - r_{AB}) + \beta(r_{Ab} - r_{AB}))(\beta + r_{aB} - r_{AB})(\alpha + r_{Ab} - r_{AB}).$$

To have $\lambda_1 < 0$, it must be that $r_{ab} < r_{AB}$. By requiring positive growth rates, $\lambda_2 < 0$. To determine the sign of λ_3 and λ_4, we will examine the restrictions obtained by requiring the populations to be positive. Note that by requiring the populations to be positive, we have that

$$\begin{aligned} \alpha + r_{Ab} - r_{AB} &> 0 \\ \beta + r_{aB} - r_{AB} &> 0 \\ -\alpha(\beta + r_{aB} - r_{AB}) + \beta(-r_{Ab} + r_{AB}) &> 0 \end{aligned}$$

which implies that δ is negative. So, $\sqrt{\gamma^2 + \delta}$ is either positive and less than $|\gamma|$, or it is imaginary. Also, by using the constraint that $\frac{\beta + r_{aB} - r_{AB}}{\beta} = x_{aB} > 0$ and using that $x_{AB} > 0$,

$$\begin{aligned} & & x_{AB} &> 0 \\ \Rightarrow & & \frac{-\alpha(\beta + r_{aB} - r_{AB}) + \beta(-r_{Ab} + r_{AB})}{\alpha\beta} &> 0 \\ \Rightarrow & & -\alpha(\beta + r_{aB} - r_{AB}) + \beta(-r_{Ab} + r_{AB}) &> 0 \\ \Rightarrow & & \beta(-r_{Ab} + r_{AB}) &> \alpha(\beta + r_{aB} - r_{AB}) \\ \Rightarrow & & (r_{AB} - r_{Ab}) &> \frac{\alpha}{\beta}(\beta + r_{aB} - r_{AB}) > 0. \end{aligned}$$

Similarly, using $\frac{\alpha + r_{Ab} - r_{AB}}{\alpha} = x_{Ab} > 0$ instead, and collecting terms in x_{AB} in the opposite fashion:

$$\begin{aligned} & & x_{AB} &> 0 \\ \Rightarrow & & \frac{-\alpha(r_{aB} - r_{AB}) + \beta(-\alpha - r_{Ab} + r_{AB})}{\alpha\beta} &> 0 \\ \Rightarrow & & -\alpha(r_{aB} - r_{AB}) + \beta(-\alpha - r_{Ab} + r_{AB}) &> 0 \\ \Rightarrow & & \alpha(r_{AB} - r_{aB}) &> -\beta(-\alpha - r_{Ab} + r_{AB}) \\ \Rightarrow & & (r_{AB} - r_{aB}) &> \frac{-\beta}{\alpha}(-\alpha - r_{Ab} + r_{AB}) \\ & & &= \frac{\beta}{\alpha}(\alpha + r_{Ab} - r_{AB}) > 0. \end{aligned}$$

Thus, $r_{AB} - r_{aB}$ and $r_{AB} - r_{Ab}$ are positive assuming that the populations are positive. This second condition can be rewritten as below, which will provide a constraint that will be useful to represent this steady state graphically.

$$(r_{AB} - r_{aB}) > \frac{\beta}{\alpha}(\alpha + r_{Ab} - r_{AB}) \Rightarrow \frac{\alpha(r_{AB} - r_{aB})}{(\alpha + r_{Ab} - r_{AB})} > \beta.$$

Using this new information, we can show

$$\begin{aligned}
\gamma &= -\alpha\beta(r_{aB} + r_{Ab} - 2r_{AB}) - \alpha(r_{aB} - r_{AB})^2 - \beta(r_{Ab} - r_{AB})^2 \\
&= \alpha\beta(r_{AB} - r_{aB}) + \alpha\beta(r_{AB} - r_{Ab}) - \alpha(r_{AB} - r_{aB})^2 - \beta(r_{AB} - r_{Ab})^2 \\
&= \alpha(r_{AB} - r_{aB})(\beta - r_{AB} + r_{aB}) + \beta(r_{AB} - r_{Ab})(\alpha - r_{AB} + r_{Ab}) > 0
\end{aligned}$$

Thus, λ_3 and λ_4 are both always negative. So, for this solution to be stable, we must have positive populations and positive growth rates and $r_{ab} < r_{AB}$.

9. The eigenvalues when

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = \left(\frac{\alpha(\beta - r_{ab} + r_{aB}) + \beta(-r_{ab} + r_{Ab})}{\alpha\beta}, \frac{r_{ab} - r_{Ab}}{\alpha}, \frac{r_{ab} - r_{aB}}{\beta}, 0\right)$$

are

$$\begin{aligned}
\lambda_1 &= -r_{ab} & \lambda_2 &= -\frac{1}{2\alpha\beta}\left(\gamma + \sqrt{\gamma^2 + \delta}\right) \\
\lambda_3 &= -\frac{1}{2\alpha\beta}\left(\gamma - \sqrt{\gamma^2 + \delta}\right) & \lambda_4 &= -r_{ab} + r_{AB}
\end{aligned}$$

with

$$\begin{aligned}
\gamma &= (\alpha + \beta)(r_{ab} - r_{aB})(r_{ab} - r_{Ab}) \\
\delta &= 4\alpha\beta(r_{ab} - r_{aB})(r_{ab} - r_{Ab})\big(\alpha(\beta - r_{ab} + r_{aB}) + \beta(-r_{ab} + r_{Ab})\big).
\end{aligned}$$

Note that γ is positive if the populations are assumed to be positive. Furthermore, δ is also positive if the populations are assumed to be positive. Then, $\gamma - \sqrt{\gamma^2 + \delta}$ is negative, so λ_3 is positive. Thus, this steady state is never stable.

A.2.4 General Cooperation Function

Consider the system below:

$$\dot{x_{ab}} = r_{ab}x_{ab} - \phi x_{ab} \tag{A.6a}$$

$$\dot{x_{Ab}} = [r_{Ab} + g_{Ab}(x_{aB}, x_{AB})]x_{Ab} - \phi x_{Ab} \tag{A.6b}$$

$$\dot{x_{aB}} = [r_{aB} + g_{aB}(x_{Ab}, x_{AB})]x_{aB} - \phi x_{aB} \tag{A.6c}$$

$$\dot{x_{AB}} = r_{AB}x_{AB} - \phi x_{AB}. \tag{A.6d}$$

The steady states and eigenvalues of the Jacobian of the system are listed below.

1. The eigenvalues when the population is of the form

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 1, 0, 0)$$

are

$$\lambda_1 = r_{ab} - r_{aB} \qquad \lambda_2 = -r_{aB}$$

$$\lambda_3 = -r_{aB} + r_{AB} \qquad \lambda_4 = -r_{aB} + r_{Ab} + g_{Ab}(1, 0).$$

Assuming positive growth rates, this state is stable if $r_{aB} > r_{ab}$, $r_{aB} > r_{AB}$ and $r_{aB} > r_{Ab} + g_{Ab}(1, 0)$.

2. The eigenvalues when the population is of the form

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 0, 1, 0)$$

are

$$\lambda_1 = r_{ab} - r_{Ab} \qquad \lambda_2 = -r_{Ab}$$

$$\lambda_3 = -r_{Ab} + r_{AB} \qquad \lambda_4 = r_{aB} - r_{Ab} + g_{aB}(1, 0).$$

Assuming positive growth rates, this state is stable if $r_{Ab} > r_{ab}$, $r_{Ab} > r_{AB}$ and $r_{Ab} > r_{aB} + g_{aB}(1, 0)$.

3. The eigenvalues when the population is of the form

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (0, 0, 0, 1)$$

are

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = -r_{AB}$$

$$\lambda_3 = r_{aB} - r_{AB} + g_{aB}(0, 1) \qquad \lambda_4 = r_{Ab} - r_{AB} + g_{Ab}(0, 1).$$

Assuming positive growth rates, this state is stable if $r_{AB} > r_{ab}$, $r_{AB} > r_{aB} + g_{aB}(0, 1)$ and $r_{AB} > r_{Ab} + g_{Ab}(0, 1)$. Then, for this steady state to be stable, the growth rates and functions must satisfy $r_{AB} > r_{ab}$, $r_{AB} > r_{aB} + g_{aB}(0, 1)$ and $r_{AB} > r_{Ab} + g_{Ab}(0, 1)$.

4. The eigenvalues when the population is of the form

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (1, 0, 0, 0)$$

are

$$\lambda_1 = -r_{ab} \qquad \lambda_2 = r_{AB} - r_{ab}$$

$$\lambda_3 = r_{aB} - r_{ab} \qquad \lambda_4 = r_{Ab} - r_{ab}.$$

Assuming positive growth rates, this state is stable if $r_{ab} > r_{AB}$, $r_{ab} > r_{aB}$ and $r_{ab} > r_{Ab}$. Then, for this steady state to be stable, the growth rates must satisfy $r_{ab} > r_{AB}$

5. For the case when the population is of the form $(0, x_{aB}, x_{Ab}, 0)$, we cannot initially calculate the eigenvalues of the Jacobian. By setting the system of differential equations to be 0, and setting $x_{ab}, x_{AB} = 0$ and using that $x_{aB} + x_{Ab} = 1$, we can determine a constraint that must be satisfied for the steady state to be stable:

$$r_{aB} + g_{aB}(x_{Ab}, 0) = r_{Ab} + g_{Ab}(x_{aB}, 0).$$

Using the constraints on the populations, it is now possible to calculate the characteristic polynomial of the Jacobian to calculate the eigenvalues. The eigenvalues for this case are

$$\lambda_1 = -r_{Ab} - g_{Ab}(x_{aB}, 0) \qquad \lambda_2 = r_{ab} - r_{Ab} - g_{Ab}(x_{aB}, 0)$$

$$\lambda_3 = -r_{Ab} + r_{AB} - g_{Ab}(x_{aB}, 0) \qquad \lambda_4 = -x_{aB}x_{Ab}\left(\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab}, 0) + \frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB}, 0)\right).$$

This state is stable if $r_{Ab} + g_{Ab}(x_{aB}, 0) > 0$, $r_{Ab} + g_{Ab}(x_{aB}, 0) > r_{Ab}$, $r_{Ab} + g_{Ab}(x_{aB}, 0) > r_{AB}$ and $x_{aB}x_{Ab}(\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab}, 0) + \frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB}, 0)) > 0$. By assuming that the populations are positive, and that the first partial derivatives are positive, we have that $\lambda_1 < 0$ and $\lambda_4 < 0$. Thus, the constraints that still must be satisfied are $r_{aB} + g_{aB}(x_{Ab}, 0) = r_{Ab} + g_{Ab}(x_{aB}, 0) > r_{ab}$ and $r_{aB} + g_{aB}(x_{Ab}, 0) = r_{Ab} + g_{Ab}(x_{aB}, 0) > r_{AB}$.

6. For the case when the population is of the form $(0, x_{aB}, 0, x_{AB})$, we can not initially

calculate the eigenvalues of the Jacobian. By setting the system of differential equations to be 0, and setting $x_{ab}, x_{Ab} = 0$ and using that $x_{aB} + x_{AB} = 1$, we can extract a constraint that must be satisfied for the steady state to be stable:

$$r_{AB} = r_{aB} + g_{aB}(0, x_{AB})$$

Using the constraints on the populations, it is now possible to calculate the characteristic polynomial of the Jacobian to calculate the eigenvalues. The eigenvalues for this case are

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = -r_{AB}$$

$$\lambda_3 = r_{Ab} - r_{AB} + g_{Ab}(x_{aB}, x_{AB}) \qquad \lambda_4 = -x_{aB}x_{AB}\frac{\partial g_{aB}}{\partial x_{AB}}(0, x_{AB}).$$

This state is stable if the product $x_{aB}x_{AB}\frac{\partial g_{aB}}{\partial x_{AB}}(0, x_{AB}) > 0$, and also $r_{AB} > r_{ab}$, $r_{AB} > 0$ and $r_{AB} > r_{Ab} + g_{Ab}(x_{aB}, x_{AB})$. By assuming that the populations are positive, and that the first partial derivatives are positive, we have that $\lambda_2 < 0$ and $\lambda_4 < 0$. Thus, the constraints that still must be satisfied are $r_{AB} > r_{ab}$ and $r_{AB} > r_{Ab} + g_{Ab}(x_{aB}, x_{AB})$.

7. For the case when the population is of the form $(0, 0, x_{Ab}, x_{AB})$, we can not initially calculate the eigenvalues of the Jacobian. By setting the system of differential equations to be 0, and setting x_{ab}, $x_{aB} = 0$ and using that $x_{Ab} + x_{AB} = 1$, we can extract a constraint on the population values that must be satisfied for the steady state to be stable:

$$r_{AB} = r_{Ab} + g_{Ab}(0, x_{AB}).$$

Using the constraints on the populations, it is now possible to find the following eigenvalues by calculating the characteristic equation of the Jacobian.

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = -r_{AB}$$

$$\lambda_3 = r_{aB} - r_{AB} + g_{aB}(x_{Ab}, x_{AB}) \qquad \lambda_4 = -x_{Ab}x_{AB}\frac{\partial g_{Ab}}{\partial x_{AB}}(0, x_{AB}).$$

This state is stable if $r_{AB} > r_{ab}$, $r_{AB} > 0$, $r_{AB} > r_{aB}+g_{aB}$ and $x_{Ab}x_{AB}\frac{\partial g_{Ab}}{\partial x_{AB}}(0, x_{AB}) > 0$. By assuming that the populations are positive, and that the first partial derivatives are positive, we have that $\lambda_2 < 0$ and $\lambda_4 < 0$. Thus, the constraints that still must be satisfied are $r_{AB} > r_{ab}$ and $r_{AB} > r_{aB} + g_{aB}(x_{Ab}, x_{AB})$.

8. For the case when the population is of the form $(0, x_{aB}, x_{Ab}, x_{AB})$, we can not initially calculate the eigenvalues of the Jacobian. By setting the system of differential equations to be 0, and setting $x_{ab} = 0$ and using $x_{aB} + x_{Ab} + x_{AB} = 1$, we can extract two constraints that must be satisfied for the steady state to be stable:

$$r_{AB} = r_{aB} + g_{aB}(x_{Ab}, x_{AB})r_{AB} = r_{Ab} + g_{Ab}(x_{aB}, x_{AB}).$$

Using the constraints on the populations, it is now possible to find the eigenvalues by calculating the characteristic polynomial. The eigenvalues are

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = -r_{AB}$$

$$\lambda_3 = \frac{1}{2}\big(\gamma - \sqrt{\gamma^2 + \delta}\,\big) \qquad \lambda_4 = \frac{1}{2}\big(\gamma + \sqrt{\gamma^2 + \delta}\,\big)$$

where

$$\begin{aligned}\gamma = & -x_{aB}x_{AB}\frac{\partial g_{aB}}{\partial x_{AB}}(x_{Ab},x_{AB}) - x_{aB}x_{Ab}\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB})\\ & - x_{aB}x_{Ab}\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB},x_{AB}) - x_{Ab}x_{AB}\frac{\partial g_{Ab}}{\partial x_{AB}}(x_{aB},x_{AB})\\ \delta = & -4x_{Ab}x_{AB}x_{aB}\Bigg(\frac{\partial g_{Ab}}{\partial x_{AB}}(x_{aB},x_{AB})\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB})\\ & + \left(\frac{\partial g_{aB}}{\partial x_{AB}}(x_{Ab},x_{AB}) - \frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB})\right)\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB},x_{AB})\Bigg).\end{aligned}$$

With positive partial derivatives and positive populations, it is clear that $\gamma < 0$. Since $\gamma < 0$, $\lambda_3 < 0$. If $\delta < 0$, then $0 < \sqrt{\gamma^2+\delta} < |\gamma|$ or $\sqrt{\gamma^2+\delta}$ is imaginary. If $0 < \sqrt{\gamma^2+\delta} < |\gamma|$, then $\lambda_4 < 0$. If $\sqrt{\gamma^2+\delta}$ is imaginary, then $\Re(\lambda_4)$ is negative. If however, $\delta > 0$, then $\sqrt{\gamma^2+\delta} > |\gamma|$, which causes $\gamma + \sqrt{\gamma^2+\delta} > \gamma + |\gamma| > 0$, which then leads to $\lambda_4 > 0$. To determine where $\delta < 0$, consider the following:

$$\begin{aligned}\delta < 0 \iff & \quad -4x_{Ab}x_{AB}x_{aB}\Bigg(\frac{\partial g_{Ab}}{\partial x_{AB}}(x_{aB},x_{AB})\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB})\\ & + \left(\frac{\partial g_{aB}}{\partial x_{AB}}(x_{Ab},x_{AB}) - \frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB})\right)\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB},x_{AB})\Bigg) < 0\\ \iff & \quad \frac{\partial g_{Ab}}{\partial x_{AB}}(x_{aB},x_{AB})\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB}) + \frac{\partial g_{aB}}{\partial x_{AB}}(x_{Ab},x_{AB})\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB},x_{AB})\\ & > \frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB})\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB},x_{AB}).\end{aligned}$$

To have stability, there are two requirements to be satisfied:

$$r_{AB} > r_{ab}$$

$$\begin{aligned}&\frac{\partial g_{Ab}}{\partial x_{AB}}(x_{aB},x_{AB})\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB}) + \frac{\partial g_{aB}}{\partial x_{AB}}(x_{Ab},x_{AB})\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB},x_{AB})\\ &\qquad > \frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB})\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB},x_{AB}).\end{aligned}$$

9. The eigenvalues where the population is of the form

$$(x_{ab}, x_{aB}, x_{Ab}, x_{AB}) = (x_{ab}, x_{aB}, x_{Ab}, 0)$$

are

$$\lambda_1 = -r_{Ab} - g_{Ab}(x_{aB}, 0) \qquad \lambda_2 = -r_{Ab} + r_{AB} - g_{Ab}(x_{aB}, 0)$$
$$\lambda_3 = \frac{1}{2}(\gamma + \sqrt{\gamma^2 + \delta}\,) \qquad \lambda_4 = \frac{1}{2}(\gamma - \sqrt{\gamma^2 + \delta}\,)$$

where

$$\gamma = -x_{aB}x_{Ab}\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab}, 0) - x_{aB}x_{Ab}\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB}, 0)$$
$$\delta = -4\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab}, 0)\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB}, 0)(-x_{aB}x_{Ab} + x_{aB}^2x_{Ab} + x_{aB}x_{Ab}^2).$$

Note that with positive first derivatives and positive populations, $\gamma < 0$. By the reasoning from the previous case, it is always the case that $\lambda_3 < 0$, and $\lambda_4 < 0$ if $\delta < 0$. Thus, this state is stable if $r_{Ab} + g_{Ab}(x_{aB}, 0) > 0, r_{Ab} + g_{Ab}(x_{aB}, 0) > r_{AB}$ and $\delta < 0$. If $\delta < 0$, then we have the following:

$$-4\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab}, 0)\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB}, 0)(-x_{aB}x_{Ab} + x_{aB}^2x_{Ab} + x_{aB}x_{Ab}^2) < 0$$
$$\Rightarrow -x_{aB}x_{Ab} + x_{aB}^2x_{Ab} + x_{aB}x_{Ab}^2 > 0$$
$$\Rightarrow x_{aB}^2x_{Ab} + x_{aB}x_{Ab}^2 > x_{aB}x_{Ab}$$
$$\Rightarrow x_{aB} + x_{Ab} > 1$$

Since we are assuming that $x_{ab}, x_{aB}, x_{Ab} > 0$, and that $x_{ab} + x_{aB} + x_{Ab} = 1$, it cannot be the case that $x_{aB} + x_{Ab} > 1$. So, this state is never stable.

A.3 Application

We will now consider the steady states, and the stability conditions of those steady states, of the system from [51] with growth rates

$$r_G = b - c_1 - c_2 \tag{A.7a}$$

$$r_1 = (1 - x_1^{n-1})b - c_1 = (1 - (1 - x_G - x_2)^{n-1})b - c_1 \tag{A.7b}$$

$$r_2 = (1 - x_2^{n-1})b - c_2 = (1 - (1 - x_G - x_1)^{n-1})b - c_2. \tag{A.7c}$$

We can apply the previous work by noticing the parallels between the system A.6d by rewriting: $r_1 = g_1(x_2, x_G) - c_1, r_2 = g_2(x_1, x_G) - c_2$, $x_1 = x_{aB}$, $x_2 = x_{Ab}$, $x_G = x_{AB}$, $x_{ab} = 0$. Then, $-c_1 = r_{aB}$, $-c_2 = r_{Ab}$, $b - c_1 - c_2 = r_{AB}$ $g_1 = g_{aB}$ and $g_2 = g_{Ab}$. Now it is straightforward to calculate the steady states.

1. $(x_1, x_2, x_G) = (1, 0, 0)$. This state will be stable if

$$\begin{aligned} \lambda_1 &= r_{ab} - r_{aB} & \lambda_2 &= -r_{aB} \\ &= -r_{aB} = c_1 < 0 & &= c_1 < 0 \\ \lambda_3 &= -r_{aB} + r_{AB} & \lambda_4 &= -r_{aB} + r_{Ab} + g_{Ab}(1, 0) \\ &= c_1 - c_1 - c_2 = c_2 < 0 & &= c_1 - c_2 + b - (1 - x_1)^{n-1} b < 0 \end{aligned}$$

Assuming c_1 and c_2 are positive, λ_1 and λ_2 show that this state is not stable.

2. $(x_1, x_2, x_G) = (0, 1, 0)$ This state will be stable if

$$\lambda_1 = r_{ab} - r_{Ab} \qquad \lambda_2 = -r_{Ab}$$

$$= c_2 < 0 \qquad = c_2 < 0$$

$$\lambda_3 = -r_{Ab} + r_{AB} \qquad \lambda_4 = r_{aB} - r_{Ab} + g_{aB}(0, 1)$$

$$= c_2 + b - c_1 - c_2 = b - c_1 < 0 \qquad = -c_1 + c_2 + b - b(1 - x_{Ab})^{n-1} < 0$$

Assuming c_1 and c_2 are positive, λ_1 and λ_2 show that this state is not stable.

3. $(x_1, x_2, x_G) = (0, 0, 1)$ This state is stable if

$$\lambda_1 = r_{ab} - r_{AB} \qquad \lambda_2 = -r_{AB}$$

$$= -b + c_1 + c_2 < 0 \qquad = -b + c_1 + c_2 < 0$$

$$\lambda_3 = r_{aB} - r_{AB} + g_{aB}(0, 1) \qquad \lambda_4 = r_{Ab} - r_{AB} + g_{Ab}(0, 1)$$

$$= -c_2 - b + c_1 + c_2 + b = c_1 < 0 \qquad = -c_1 - b + c_1 + c_2 + b = c_2 < 0$$

Assuming c_1 and c_2 are positive, λ_3 and λ_4 show that this state is not stable.

4. $(x_1, x_2, x_G) = (x_1, x_2, 0)$ This state is stable if $r_{aB} + g_{aB}(x_{Ab}, 0) > r_{AB}$ and $r_{Ab} + g_{Ab}(x_{aB}, 0) > r_{AB}$ and the populations are positive. For the inequalities, we want to show:

$$b - c_A - c_B < b - c_A - b(1 - x_{Ab})^{n-1}$$

$$b - c_A - c_B < b - c_B - b(1 - x_{aB})^{n-1}$$

these become:

$$-c_B < b - b(1 - x_{Ab})^{n-1}$$

$$-c_A < b - b(1 - x_{aB})^{n-1}$$

and we have:

$$b(1-(1-x_{Ab})^{n-1}>0, \text{ and } -c_B<0$$
$$b(1-(1-x_{Ab})^{n-1}>0, \text{ and } -c_A<0$$

so the inequalities hold. Furthermore, the populations must satisfy the following equality:

$$\begin{aligned} x_{Ab}^{n-1} &= \frac{c_B-c_A}{b}+x_{aB}^{n-1} \\ &= \frac{c_B-c_A}{b}+(1-x_{Ab})^{n-1} \end{aligned}$$

which can be used to shown that the populations are positive. Based on initial conditions, $x_{Ab}\in[0,1]$, so $(1-x_{Ab})^{n-1}\in[0,1]$. If $c_B-c_A>0$ then $x_{Ab}^{n-1}>(1-x_{Ab})^{n-1}\Rightarrow x_{Ab}>1-x_{Ab}\Rightarrow 2x_{Ab}>1\Rightarrow x_{Ab}>1/2$. Furthermore, $x_{Ab}=1$ is only a solution to this equality if $c_B-c_A=b$, which is assumed to not be the case. So this state is stable.

5. $(x_1,x_2,x_G)=(x_1,0,x_G)$ For this state to be stable, the eigenvalue $\lambda_2=-r_{AB}=-b+c_1+c_2<0$, which is assumed to not be the case. So this state is not stable.

6. $(x_1,x_2,x_G)=(x_1,x_2,x_G)=(c_B^{1/(-1+n)},c_A^{1/(-1+n)},1-c_A^{1/(-1+n)}-c_B^{1/(-1+n)})$ For this state to be stable, we must have positive populations and

$$\begin{aligned} &\frac{\partial g_{Ab}}{\partial x_{AB}}(x_{aB},x_{AB})\frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB})+\frac{\partial g_{aB}}{\partial x_{AB}}(x_{Ab},x_{AB})\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB},x_{AB}) \\ &\qquad > \frac{\partial g_{aB}}{\partial x_{Ab}}(x_{Ab},x_{AB})\frac{\partial g_{Ab}}{\partial x_{aB}}(x_{aB},x_{AB}) \end{aligned}$$

Note that for any c_A and c_B, there is an n such that $c_A^{1/(-1+n)}+c_B^{1/(-1+n)}>1$ and so the state is then not stable based on our initial conditions for large enough n.

Appendix B

Supplementary Material to Chapter 3

B.1 Variations of Functional Forms

The function expression for the payoff function that was used in most of the illustrations has the form:

$$\begin{aligned}P(\alpha, \beta, a, b) = &S_{h,m}(\alpha + a)S_{h,m}(\beta + b)\\&- \delta(\alpha + \beta) - \gamma S_{j,n}(\alpha)S_{j,n}(\beta),\end{aligned}$$

where (α, β) is the ordered pair trait value of the mutant, and (a, b) is the ordered pair trait value of the environment. We furthermore assumed that δ, γ, h, j, were all positive with m and n greater than zero, and the sigmoid function is the Hill function:

$$S_{h,m}(x) = \frac{x^m}{x^m + h^m}.$$

In this section, we examine other functional forms compatible with equation (3.1) of the main text, as well as an example that is incompatible with it.

B.1.1 Different Sigmoid Function for Benefit Term

We can use different sigmoid functions for the benefit term and still see branching. For example, we implemented the arctangent function:

$$P(\alpha, \beta, a, b) = \arctan[m(a + \alpha - h)] \arctan[m(b + \beta - h)] - \delta(\alpha + \beta) - \gamma S_{j,n}(\alpha) S_{j,n}(\beta), \tag{B.1}$$

for an individual with trait value (a, b) in an environment of (α, β). Figure B.1 demonstrates that branching can still be observed in this case.

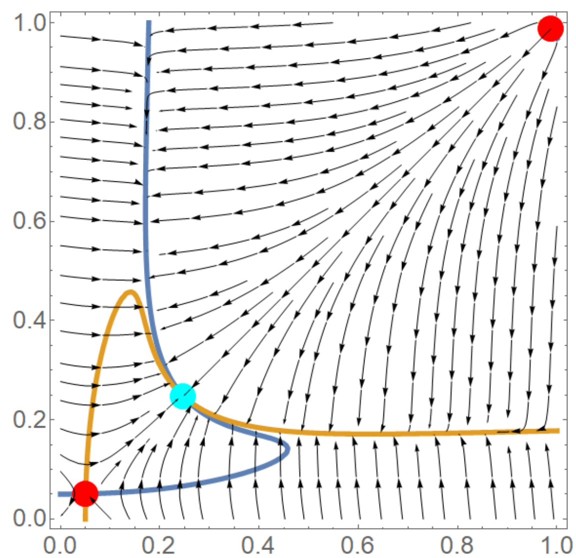


Figure B.1: Model (B.1): Singular strategies, evolutionary isoclines and streamplot of the selection gradients for parameters $\delta = .1$, $\gamma = .75$, $h = 0$, $m = 1$, $j = .4$, $n = 3$. The stability of the singular strategies are color-coded per table 3.2 in the main text.

B.1.2 A Non-saturating Function for Benefit Term

If we do not use a saturating function, we may lose the steady state that leads to branching. For instance, if we use the pure product ab rather than the product of two saturating functions, that is, set

$$P(\alpha, \beta, a, b) = (a + \alpha - h)(b + \beta - h) - \delta(\alpha + \beta) - \gamma S_{j,n}(\alpha) S_{j,n}(\beta), \tag{B.2}$$

the trait values can increase without bound given an initial condition larger than the unstable strategy shown in figure B.2.

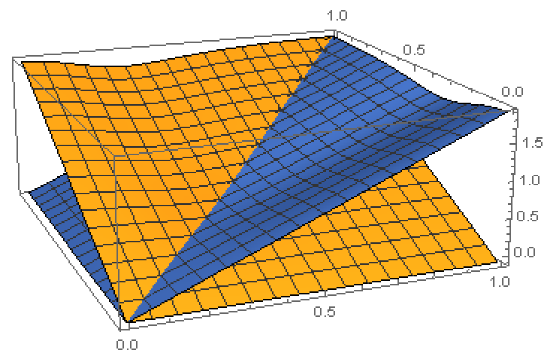


Figure B.2: Model (B.2): the first derivatives of the payoff functions. The singular strategies occur where the surfaces are both zero. Parameters are $\delta = .1$, $\gamma = .75$, $j = .4$, $n = 3$ (note that h and m are not present in this list, since the benefit is not defined in terms of a sigmoid function). There is a singular strategy at (.05, .05), which is unstable. This image shows that the first derivatives will not be zero as the traits increase, indicating that there will not be another singular strategy.

B.1.3 Quadratic Individual Task Cost

Next, we will alter this model by changing the individual task cost term so that α and β are each squared. This gives us the new payoff function

$$P(\alpha, \beta, a, b) = S_{h,m}(\alpha + a) S_{h,m}(\beta + b) - \delta(\alpha^2 + \beta^2) - \gamma S_{j,n}(\alpha) S_{j,n}(\beta); \tag{B.3}$$

we will keep all the assumptions about the parameters as we did previously. In this case, we still expect branching for a subset of parameters, as shown in figure B.3.

The system with a quadratic individual cost term has some interesting properties; namely, it

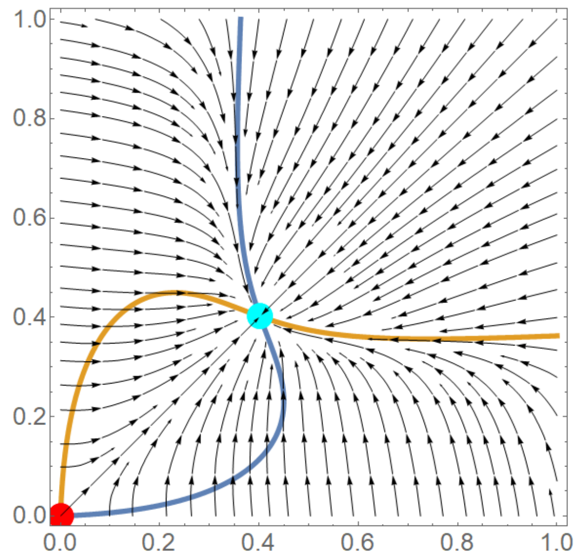


Figure B.3: Model (B.3): singular strategies, evolutionary isoclines and streamplot of the selection gradients for parameters $\delta = .1$, $\gamma = .1$, $h = .6$, $j = .4$, $m = 1$, $n = 3$ with $\delta(a^2+b^2)$ rather than $\delta(a+b)$ in the cost term. The stability of the singular strategies are color-coded per table 3.2 in the main text.

allows the origin $(0, 0)$ to be an unstable equilibrium. To study this system, we will examine the selection gradients of the fitness function ($F(\alpha, \beta, a, b) = P(\alpha, \beta, a, b) - P(a, b, a, b)$), which we will, as before, call F_α and F_β.

$$F_\alpha(a, b) = S'_{h,m}(2a)S_{h,m}(2b) - 2\delta a - \gamma S'_{j,n}(a)S_{j,n}(b)$$
$$F_\beta(a, b) = S_{h,m}(2a)S'_{h,m}(2b) - 2\delta b - \gamma S_{j,n}(a)S'_{j,n}(b).$$

We are assuming that the Hill-type functions are zero at zero, so we can see that $F_\alpha(0, 0) = F_\beta(0, 0) = 0$, indicating that the origin will be a singular strategy. To determine the stability, we must also examine the eigenvalues of the Jacobian of these selection gradients.

$$J = \begin{bmatrix} 2S''_{h,m}(2a)S_{h,m}(2b) - 2\delta - \gamma S''_{j,n}(a)S_{j,n}(b) & 2S'_{h,m}(2a)S'_{h,m}(2b) - \gamma S'_{j,n}(a)S'_{j,n}(b) \\ 2S'_{h,m}(2a)S'_{h,m}(2b) - \gamma S'_{j,n}(a)S'_{j,n}(b) & 2S_{h,m}(2a)S''_{h,m}(2b) - 2\delta - \gamma S_{j,n}(a)S''_{j,n}(b) \end{bmatrix}.$$

At this point we will assume a particular form of the Hill-type functions. We will assume that

$$S_{h,m}(x) = \frac{x^m}{x^m + h^m}$$

where m is a positive whole number. Then, if $m > 1$:

$$S'_{h,m}(x) = \frac{mx^{m-1}h^m}{(x^m + h^m)^2}$$
$$S''_{h,m}(x) = \frac{(m-1)mh^m x^{m-2}}{(x^m + h^m)^2} - \frac{2m^2 h^m x^{2m-2}}{(x^m + h^m)^3}.$$

And so, if both $m > 1$ and $j > 1$ for our Hill functions, we have at the origin the Jacobian is

$$J = \begin{bmatrix} -2\delta & 0 \\ 0 & -2\delta \end{bmatrix}.$$

This leaves us with eigenvalues -2δ and -2δ, both negative, indicating that the singular strategy at the origin is convergent stable (to determine evolutionary stability we would have to examine the Hessian). This will prevent the emergence of cooperative behavior, but does not invalidate our claims of branching when cooperative behavior is already in place.

We will use the same form for the Hill-type functions:

$$S_{h,m}(x) = \frac{x^m}{x^m + h^m}.$$

Then, if $m = 1$, we have

$$S'_{h,m}(x) = \frac{h}{(x+h)^2}$$
$$S''_{h,m}(x) = -2h(x+h)^{-3}.$$

If $m = 1$ in our payoff function, the Jacobian at the origin becomes

$$J = \begin{bmatrix} -2\delta & \frac{1}{h^2} \\ \frac{1}{h^2} & -2\delta \end{bmatrix}$$

which has eigenvalues $-2\delta \pm \frac{1}{h^2}$. When one of these eigenvalues are positive ($2\delta < \frac{1}{h^2}$), then the origin will not be convergent stable.

Similarly, if $m = 1$ and $n = 1$ in our payoff function, the Jacobian at the origin becomes

$$J = \begin{bmatrix} -2\delta & \frac{1}{h^2} - \frac{1}{j^2} \\ \frac{1}{h^2} - \frac{1}{j^2} & -2\delta \end{bmatrix}$$

which has eigenvalues $-2\delta \pm (\frac{1}{h^2} - \frac{1}{j^2})$. When one of these eigenvalues are positive ($2\delta + \frac{1}{j^2} < \frac{1}{h^2}$), then the origin will not be convergent stable.

Also, if $n = 1$, but $m > 1$, the Jacobian at the origin becomes

$$J = \begin{bmatrix} -2\delta & -\frac{1}{j^2} \\ -\frac{1}{j^2} & -2\delta \end{bmatrix}$$

which has eigenvalues $-2\delta \pm \frac{1}{j^2}$. When one of these eigenvalues are positive ($2\delta < \frac{1}{j^2}$), then the origin will not be convergent stable.

To illustrate this analysis, figure B.4 presents evolutionary dynamics of this system starting near the origin. We can see that first, the monomorphic population follows a trajectory away from $(0, 0)$ towards an equilibrium characterized by division of labor. At this point, branching occurs and the two populations of specialists evolve.

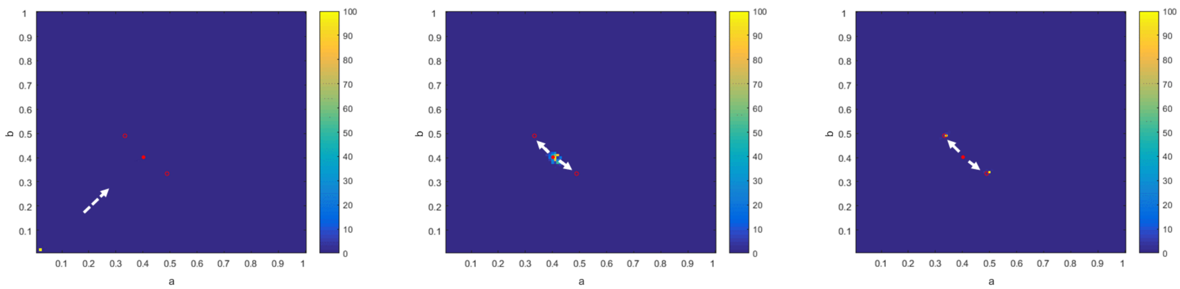


Figure B.4: Model (B.3): heat maps of trait values for quadratic individual task costs. Parameters are $\delta = .1$, $\gamma = .1$, $h = .6$, $j = .4$, $m = 1$, $n = 3$.

B.1.4 Different Form for the Additional Cost for Performing Both Tasks

We can also change the form of the second cost term in equation (3.1) of the main text. For example, we can assume that the total cost has the form $\delta(a + b) + \gamma ab$:

$$P(\alpha, \beta, a, b) = S_{h,m}(\alpha + a)S_{h,m}(\beta + b) - \delta(\alpha + \beta) - \gamma ab. \tag{B.4}$$

In this case we can again observe branching, as shown in figure B.5. A condition for branching in this case is a relatively small value for δ relative to γ. We also observe branching for the case when $\delta = 0$. For biological reasons, we would like to have the cost associated with δ because it can represent factors such as time investment or maintenance of structures to perform a task.

Similar to the work done in the main paper, we investigated the parameter space to find out where we may expect branching to occur, see figure B.6. In this figure, $m = 1$ and $h = .5$ are fixed, and parameters δ and γ are varied.

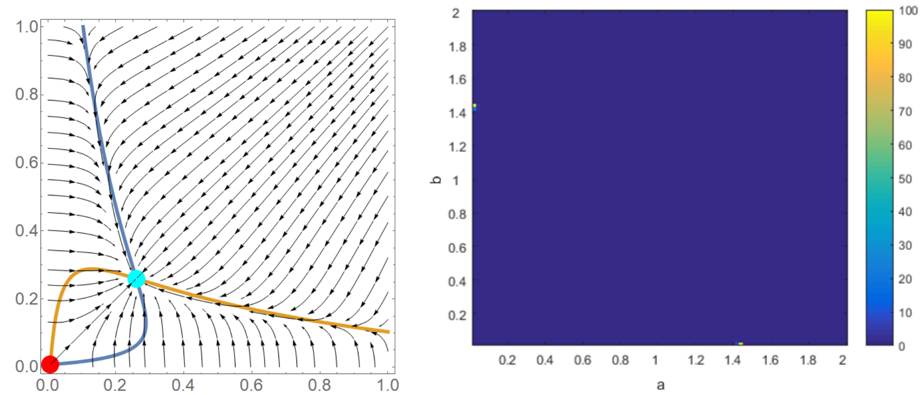


Figure B.5: Dynamics of model (B.4). Left: Singular strategies, evolutionary isoclines and streamplot of the selection gradients for parameters $\delta = .05$, $\gamma = .75$, $h = .5$, $m = 1$. Note that j and n are not present in this list, as we are not using a sigmoid function for the additional cost of performing both tasks. The stability of the singular strategies are color-coded per table 3.2 in the main text. Right: Final simulation heat map of trait values, after 10^6 births. The simulation shows convergence to two cooperative populations located at the horizontal and the vertical axes. Initial trait values were both .45.

B.1.5 A Simplified Model

Finally, we present analysis of a simple model which used multiplication rather than products of Hill functions for both benefits and costs:

$$P(a, b, \alpha, \beta) = (a + \alpha)(b + \beta) - \delta(a + b) - \gamma ab. \tag{B.5}$$

The selection gradients are

$$\begin{aligned}\frac{\partial P}{\partial a}(a, b, \alpha, \beta)|_{a=\alpha, b=\beta} &= 2\beta - \delta - \gamma b \\ \frac{\partial P}{\partial b}(a, b, \alpha, \beta)|_{a=\alpha, b=\beta} &= 2\alpha - \delta - \gamma \alpha.\end{aligned}$$

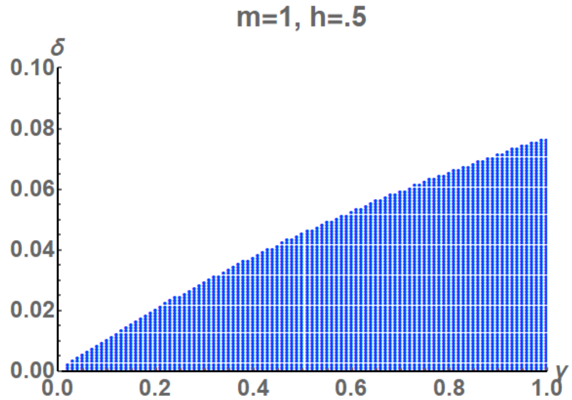


Figure B.6: The system with an alternative form for the additional cost for performing both tasks, model (B.4). The points indicate parameter values for which branching is expected to occur. All branching is expected to occur along the (-1,1) direction. For this image, $m = 1$ and $h = .5$ while δ is varied along the vertical axis and γ along the horizontal axis.

We can then find the singular strategies: $\beta = \frac{\delta}{2-\gamma}$ and $a = \frac{\delta}{2-\gamma}$. The Jacobian of the selection gradients determines the stability of the strategies found from the selection gradients:

$$J = \begin{bmatrix} 0 & 2-\gamma \\ 2-\gamma & 0 \end{bmatrix}.$$

The eigenvalues are $2 - \gamma$ and $-2 + \gamma$. Clearly, there are no parameters which allow for any steady states to be convergent stable, which is when both eigenvalues are negative. This means that the steady strategy is not convergent stable, see figure B.7. Examining the Hessian, we see that

$$H = \begin{bmatrix} 0 & 1-\gamma \\ 1-\gamma & 0 \end{bmatrix}$$

The Hessian has eigenvalues $1 - \gamma$ and $-1 + \gamma$. To have branching, we must have a positive and negative eigenvalue. We require that γ be positive since it is part of the cost. Then, for $0 < \gamma < 1$, the first eigenvalue is positive, and the second is negative, and the other way around for $\gamma > 1$. There is however no stable strategy to converge to, and therefore the population will not branch, as confirmed through simulations.

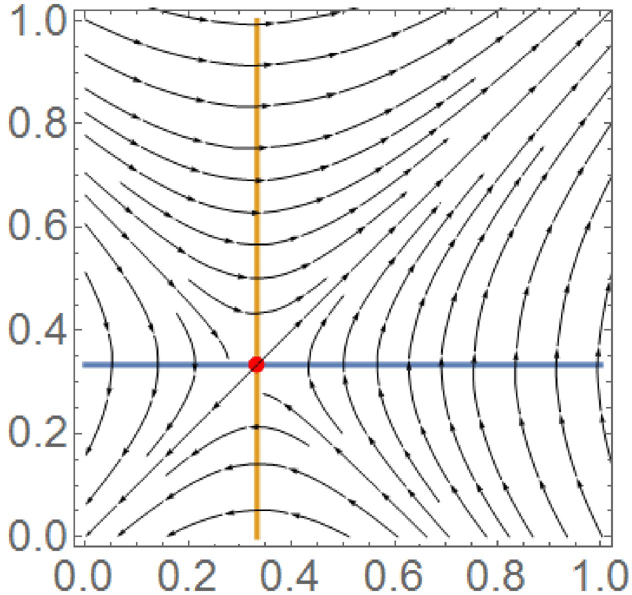


Figure B.7: Phase plane diagram for payoff (B.5) with $\delta = .5$, $\gamma = .5$.

B.2 Single Trait Model

It is interesting to examine the case in which only a single trait is evolved, as the results differ from those seen in the two trait case. We will consider the basic model for the payoff discussed in the main text, and investigate the case where one of the traits is held fixed. This would be equivalent to the two trait problem, but with trait values (a, b) for the mutant and (α, b) for the resident, with b remaining constant. We have

$$P(a, \alpha) = S_{h,m}(a + \alpha)S_{h,m}(2b) - \delta(a + b) - \gamma S_{j,n}(a)S_{j,n}(b), \tag{B.6}$$

where b is the fixed value of the other trait. Then, the selection gradient is

$$\begin{aligned} F_a(\alpha) &= \frac{\partial P(a, \alpha)}{\partial a}|_{a=\alpha} \\ &= S'_{h,m}(a + \alpha)S_{h,m}(2b) - \delta - \gamma S'_{j,n}(a)S_{j,n}(b)|_{a=\alpha} \\ &= S'_{h,m}(2\alpha)S_{h,m}(2b) - \delta - \gamma S'_{j,n}(\alpha)S_{j,n}(b). \end{aligned}$$

If we set $F_a(\alpha) = 0$, we could determine the singular strategy. We will denote the singular strategy by a^*. Then, to determine the convergent stability, we would examine the first derivative of the selection gradient, and evaluate it at the singular strategy a^*:

$$\frac{\partial F_a(\alpha)}{\partial \alpha}|_{\alpha=\alpha^*} = 2S''_{h,m}(2\alpha^*)S_{h,m}(2b) - \gamma S''_{j,n}(\alpha^*)S_{j,n}(b). \tag{B.7}$$

If this statement is positive, then the singular strategy is not convergent stable; if it is negative, then the singular strategy is convergent stable. To determine the evolutionary stability, we would need to examine the second derivative of the payoff function with respect to the mutant individual, evaluated at the singular strategy:

$$\frac{\partial^2 P(a,\alpha)}{\partial a^2}|_{a=\alpha=\alpha^*} = S''_{h,m}(2\alpha^*)S_{h,m}(2b) - \gamma S''_{j,n}(\alpha^*)S_{j,n}(b). \tag{B.8}$$

Now, we would like to be able to say something about the convergent and evolutionary stability. Fortunately, we have some properties of the Sigmoid functions we can utilize. For example, we are assuming positive sigmoid functions. In addition, we can note that second derivatives are positive for inputs less than h and j respectively, and negative for inputs larger than h and j. However, without knowing the relationship between h and j, or where α^* is relative to them, we cannot say which case we are in. So, we will exhaustively examine the scenarios, under the assumption that the singular strategy α^* is at least convergent stable:

1. If $S''_{h,m}(2\alpha^*) > 0$ and $S''_{j,n}(\alpha^*) > 0$ then if equation B.7 is negative (for α^* to be a convergent stable strategy) then equation B.8 is also negative, implying that the singular strategy is evolutionarily stable as well.

2. If $S''_{h,m}(2\alpha^*) > 0$ and $S''_{j,n}(\alpha^*) < 0$ then equation B.7 cannot be negative, since equation B.7 will be the sum of two positive terms.

3. If $S''_{h,m}(2\alpha^*) < 0$ and $S''_{j,n}(\alpha^*) > 0$ then equation B.7 is negative and so is equation B.8. Note that this differs from the first case, in which we are not guaranteed that the singular strategy will be convergent stable.

4. If $S''_{h,m}(2\alpha^*) < 0$ and $S''_{j,n}(\alpha^*) < 0$ then we cannot draw any conclusions. Even if equation B.7 is negative, we still cannot say anything about the evolutionary stability.

So, much of the analysis will depend on the particular values that we choose to examine.

B.2.1 Particular Values

We will examine this system now for a particular set of parameters to compare it to the two trait system. In particular, we will use a set of parameters that is known to have a branching point: $\delta = .1,\ \gamma = .3,\ h = .75,\ j = .5,\ m = 6,\ n = 6$. We will also set $b = .51508$, the point at which the branching occurs in the full, two trait model. In this case the singular strategy is found by solving

$$\frac{\partial P(a,\alpha)}{\partial a}\Big|_{a=\alpha} = 0$$

which gives us that the singular strategies are $\alpha^* = .163201$ or $.51508$. The stability of these strategies is determined by

$$\left(\frac{\partial}{\partial \alpha}\frac{\partial P(a,\alpha)}{\partial a}\Big|_{a=\alpha}\right)\Big|_{\alpha=\alpha^*}.$$

At $\alpha^* = .163201$, this second derivative is 3.0114, indicating instability of the singular strategy. At $\alpha^* = .51508$, this second derivative is -4.64037, indicating that it is convergent stable. The trait value will branch upon reaching the singular strategy if

$$\frac{\partial^2 P(a, \alpha)}{\partial a^2}\Big|_{\alpha=a=\alpha^*} > 0.$$

At $\alpha^* = .51508$, it evaluates to -1.61774, so branching is not expected upon reaching the singular strategy.

So, although branching is expected when there are two traits, it is not expected when there is only one. Intuitively this can be explained in the following way: if a mutation has a higher trait value a than the population, it will incur additional cost without much improvement to the benefit it receives. So, the population is not invadible from above, therefore branching cannot occur.

B.3 Different Numbers of Interactions

In the main text, the fitness was calculated as a total payoff received by an individual from interacting with all the rest of the individuals in the population. This is not the only possible formulation, and it is interesting to investigate how results change with the number of interactions included in the calculation of the fitness values. A detailed study of this issue is a work in progress, but here we report the main trends relevant for the present study.

We have performed simulations using one interaction, ten interactions, and all possible interactions, and observed that the main message of the paper is independent of this: branching behavior is still observed for a range of parameters in systems described by equation (3.1) of the main text. The quantitative behavior can however be different, and in the case where the Hessian has two positive eigenvalues, we may even see some qualitatively different behavior.

In figure B.8 we present an example of directional branching (branching is expected along

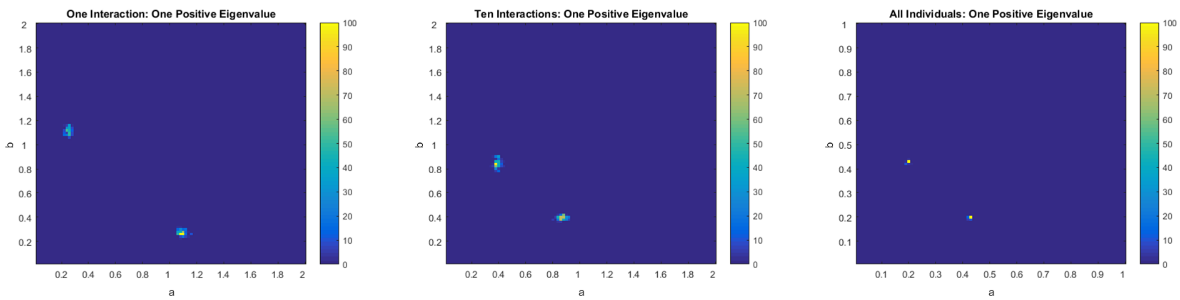


Figure B.8: Heat maps of trait values when one interaction, ten interactions, and all individuals are used to determine the fitness of individuals in the simulation. Parameters are $\delta = .1$, $\gamma = .5$, $h = .75$, $j = .5$, $m = 6$, $n = 6$. For these parameters, the Hessian indicates branching populations will converge to points along the $(-1, 1)$ vector after converging to the stable singular strategy.

the $(-1, 1)$ vector), and study the effect of changing of the number of interactions. We notice that the behavior in all three cases is qualitatively similar, but the the trait values to which the sub-populations converge are somewhat different.

A more interesting case-study is presented in figure B.9, where disrupting branching is expected (the Hessian has two positive eigenvalues). Although branching is observed for all three cases presented (one interaction, ten interactions, and all individuals), there are significant differences in the post-branching behavior. When only one interaction is used for determining the winner in the simulation, there are only two trait profiles present. Increasing to ten interactions, or using all the individuals in the calculation, we observe that there are three trait profiles that persist. This shows that the number of interactions used in the simulation can qualitatively affect the results of the simulation.

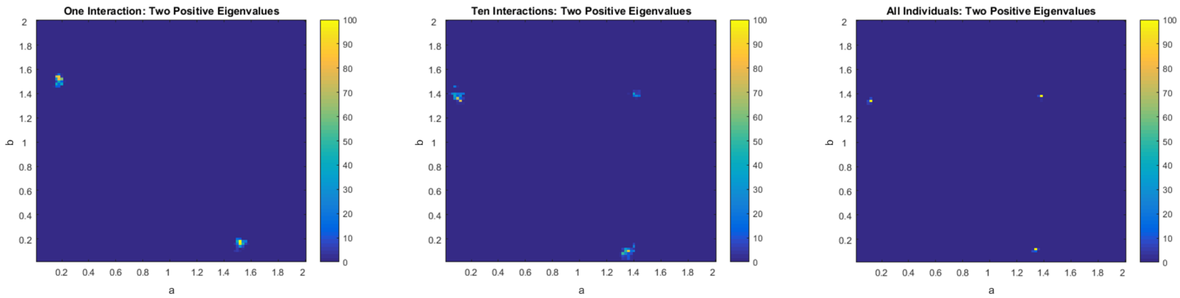


Figure B.9: Heat maps of trait values when one interaction, ten interactions, and all individuals are used to determine the winning individual in the simulation. Parameters are $\delta = .1$, $\gamma = .3$, $h = .8$, $j = .4$, $m = 3$, $n = 3$. For these parameters, the Hessian has two positive eigenvalues.

B.4 An Alternative Approach: Three Population Profiles

We would like to compare the approach in the paper with an alternative approach, where deterministic dynamics of several coexisting populations is studied. Let us suppose that we would like to find out whether cooperation between two populations is preferable to a single population of individuals that is capable of performing all the tasks. It might be tempting to consider only three types of individuals: two specialized types, (1,0) and (0,1) with frequencies x_1 and x_2 respectively, and an individual which performs both tasks, $(1,1)$ with frequency x_G. Using our payoff function, and assuming a population of three types, the average payoffs can be easily found:

$$
\begin{aligned}
P_{(1,0)} &= x_1 S_{h,m}(2) S_{h,m}(0) + x_2 S_{h,m}(1) S_{h,m}(1) + x_G S_{h,m}(2) S_{h,m}(1) - \delta \\
P_{(0,1)} &= x_1 S_{h,m}(1) S_{h,m}(1) + x_2 S_{h,m}(0) S_{h,m}(2) + x_G S_{h,m}(1) S_{h,m}(2) - \delta \\
P_{(1,1)} &= x_1 S_{h,m}(2) S_{h,m}(1) + x_2 S_{h,m}(2) S_{h,m}(2) + x_G S_{h,m}(2) S_{h,m}(2) - 2\delta - \gamma S_{j,n}(1) S_{j,n}(1) \\
\bar{P} &= x_1 P_{(1,0)} + x_2 P_{(0,1)} + x_G P_{(1,1)}
\end{aligned}
$$

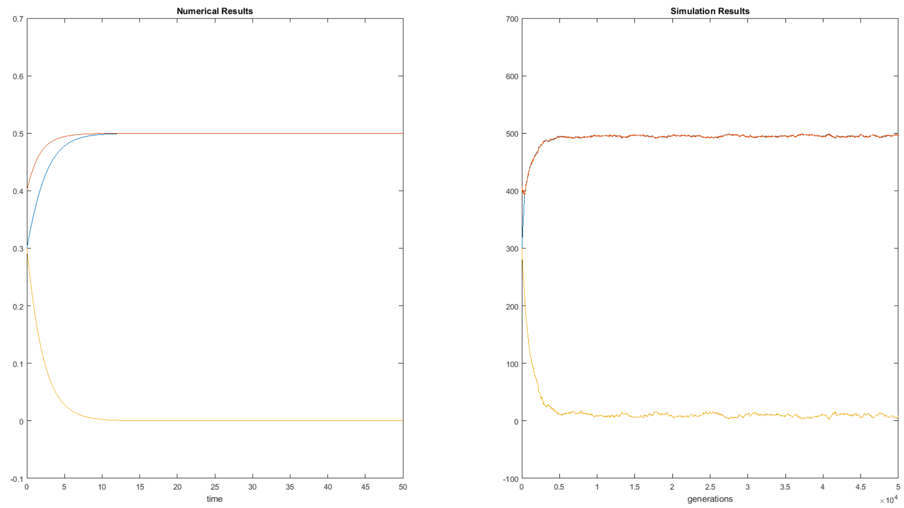


Figure B.10: (Left) Numerical Solution to system B.9. (Right) Simulation population values for three population profiles

Then, the differential equation system to describe the frequencies is as follows:

$$x_1' = (P_{(1,0)} - \bar{P})x_1 \tag{B.9a}$$

$$x_2' = (P_{(0,1)} - \bar{P})x_2 \tag{B.9b}$$

$$x_G' = (P_{(1,1)} - \bar{P})x_G \tag{B.9c}$$

One could then proceed with the usual stability analysis to find out under what parameter combination an equilibrium involving nontrivial populations x_1 and x_2 is stable, and when the equilibrium with $x_1 = x_2 = 0$ and $x_G > 0$ to be stable. For example, under parameter choices: $\delta = .5, \gamma = .5, h = .5, j = .5, m = 6$ and $n = 6$ we expect that the equilibrium with $x_1 = x_2 > 0$, $x_G = 0$ to be stable. This case is illustrated in figure B.10, left panel.

On the other hand, the adaptive dynamics for the same parameter values predicts the population to converge to the convergent and evolutionarily stable strategy, (0.350845, 0.350845). This is illustrated by figure B.11. We can see a contradiction between the two theories, which is easily resolved if the note the following.

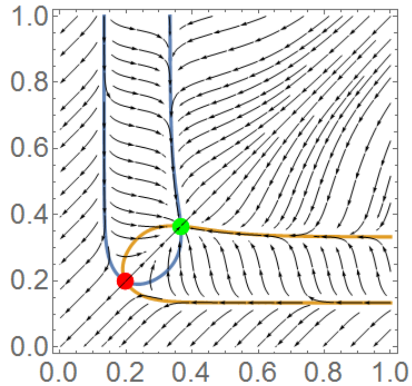


Figure B.11: Evolutionary isoclines and singular strategies for parameters $\delta = .5$, $\gamma = .5$, $h = .5$, $j = .5$, $m = 6$, $n = 6$.

In system B.9, it is simply not possible to have an individual with trait profile (0.365713, 0.365713), so this cannot be a solution to the system. Instead, we see coexistence of the two specialized populations, as shown in figure B.10 (left panel). If however we restrict the adaptive dynamics simulation to only allow for the three populations included in system B.9, then we observe the behavior predicted by the numerical solution to system B.9. In figure B.10 (right), the simulation was restricted to only have three types of individuals: specialized individuals of type $(1, 0)$ and $(0, 1)$, and individuals of type $(1, 1)$ which perform both tasks. The simulation was run with a total population of 1000 and initial sub population sizes of 300, 400 and 300 for the two specialized populations and the population which performs both tasks, respectively. The number of each type was tracked over 50000 updates to the population. The resulting populations are shown in figure B.10 (right panel).

Unlike in the numerical solution to system B.9, in the simulation, the population which performs both tasks is not eliminated. This is due to the presence of mutations, which allows the population to be maintained. The system B.9 does not have mutations accounted for, so the individuals which perform both tasks are not able to maintain any population.

Appendix C

Supplementary Material for Chapter 4

C.1 On the Behavior of the Probabilities of Fixation

Figure 4.5, shows the probability of fixation for the different cases that arise when the deterministic criterion is used. There is some interesting behavior in this image. For example, a_4 decays quickly near 440. From equation (4.36), the probability of absorption formula, there is only a 6% chance of fixation by a mutant when there are 440 cheaters. By performing a Markov process simulation in MATLAB, we get a similar probability, so we have reason to believe this is correct, and not a numerical error. By examining how $\prod q(i)/p(i)$ changes (part of equation (4.36)) we can see why the population decays so quickly. For a_4 these products are shown in figure C.1.

In figure 4.5, we see that a_2 follows with a_4 closely, but it does not experience the sudden drop, so it might seem reasonable to examine how a_2 is calculated. However, the ratio between the probability of increase and decrease for a_2 does not depend on i, while the ratios for a_4 do, so for a fixed c, $q(i)/p(i)$ is a fixed value, and this value varies slowly as the number of cheaters increase. Rather than look at the behavior of $q(i)/p(i)$ for a_2, we will

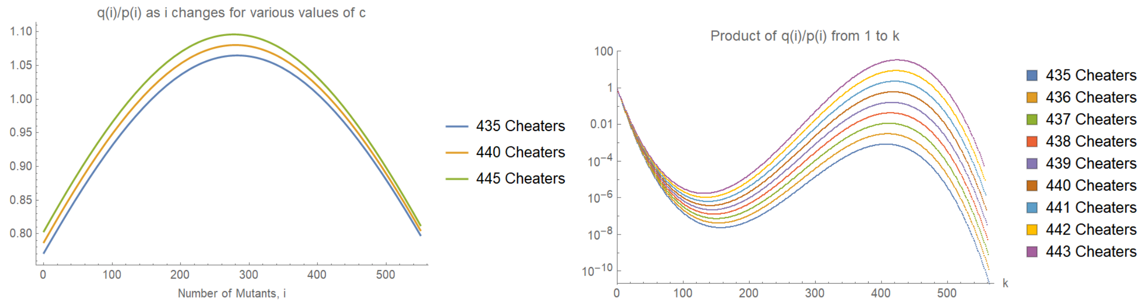


Figure C.1: To explain the sudden drop of a_4, we examined $q(i)/p(i)$ (left) and $\prod_{i=1}^{k} q(i)/p(i)$ (right), both part of equation (4.36), which we used to calculate a_4. We calculated this for several population sizes of cheaters in the region in which we see the sudden drop in the value of a_4. For the right image, note that the vertical axis has a log scale to better show the change in the product.

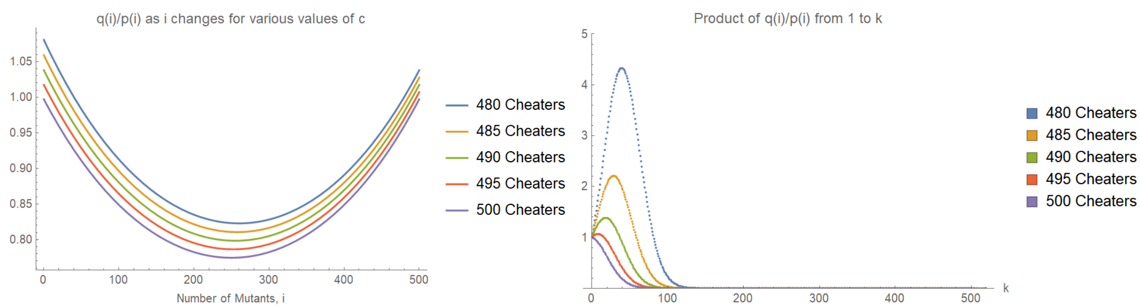


Figure C.2: Here we examine $q(i)/p(i)$ and $\prod_{i=1}^{k} q(i)/p(i)$, for $q(i)$ and $p(i)$ in the a_5 case. We calculated this for several population sizes of cheaters in the region in which we see a_5 become non-zero. For the right image, note that the vertical axis has a linear scale, as opposed to figure C.1(right).

instead examine it for a_5, where the ratio for the probability of increase and decrease for a_5 does depend on i. Unlike a_4, a_5 does not experience a sudden change, making it a good candidate to examine. We can see the change as the number of cheaters increases in figure C.2. In this figure the scale is linear instead of the logarithmic scale used in figure C.1, and the change happens so slowly that we showed curves as the number of cheaters increased by fiver rather than show curves as the number of cheaters increases by one.

C.2 Comparison to Simulations and Dangers of Large Mutation Rates

We will now compare simulations to the constructed ODEs for the deterministic criterion. We will use the following payoff function to describe the payoff that an individual with trait x receives when interacting with an individual with trait y:

$$F(x,y) = S_h(x+y)^2 - (2\delta x + \gamma S_g(x)^2),$$

where δ and γ are constants, and $S_h(x) = \frac{x^m}{x^m+h^m}$ and $S_g(x) = \frac{x^m}{x^m+g^m}$, which are Hill functions that have inverse width m, and the halfway point of the increase occurs at h and g respectively.

The requirement that mutations happen rarely and that the standard deviation of the mutations is small is more important for some parameter sets than it is for others. In figure C.3 (left), we can see that although initially the behavior is somewhat rugged, it does settle to the predicted value from the ODE. Lowering the probability of mutation and the standard deviation from the parent's trait for the mutant's trait can fix these discrepancies, as we can see in figure C.3 (right). These discrepancies are due to many mutations being present at once, causing the population to deviate from the assumed monomorphic (or dimorphic, when a single mutation is present) profile.

Let us consider the case where $\delta = .1$, $\gamma = .5$, $h = .5$, $g = .2$, $m = 6$, with 450 cheaters in a total population of 1000. By examining the function for the right hand side of the ODE, graphed in figure C.4, we can see that there is a region where the right hand side of the ODE will be equal to zero between approximately .47 and .53. In addition, we expect the traits to increase to this region if we start just below it, and decrease to it if we start just above it.

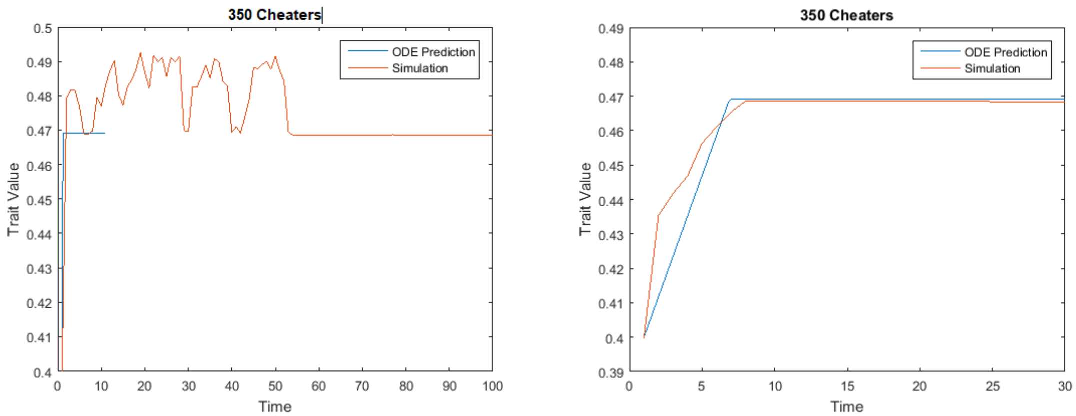


Figure C.3: Comparison of the ODE and stochastic simulation behavior when there are 350 cheaters and 650 non-cheaters. The deterministic criterion was used in the simulation and the ODE is the one derived for that criterion. Parameters are $\delta = .1$, $\gamma = .5$, $h = .5$, $g = .2$, $m = 6$. We can see that there is consistency between the long term behavior of the simulation and the ODE. Left: standard deviation of mutant from parent trait is .01, probability of mutation is .005 and the time scale for the ODE has been adjusted to match the simulation. Right: standard deviation of mutant from parent trait is .001, probability of mutation is .0005.

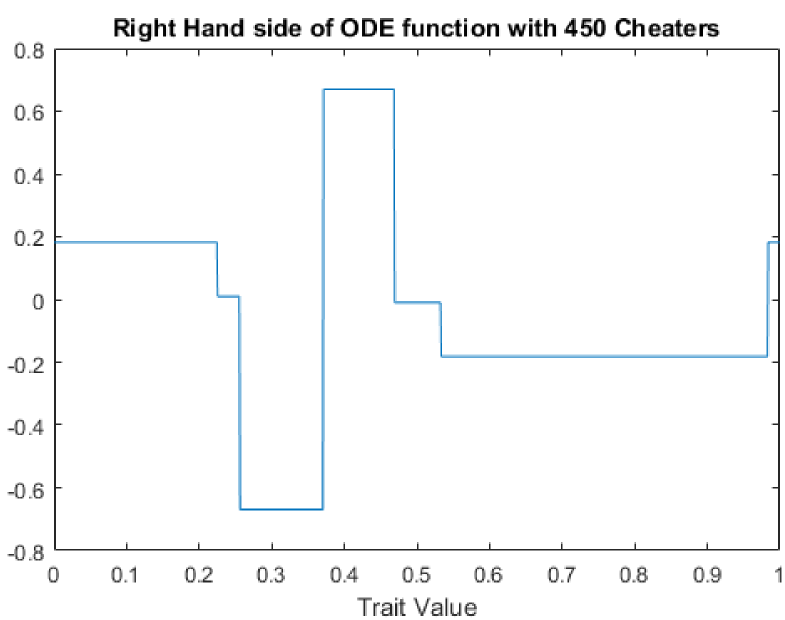


Figure C.4: The function for the right hand side of the ODE as a function of the trait value of the resident when $\delta = .1$, $\gamma = .5$, $h = .5$, $g = .2$, $m = 6$, and there are 450 cheaters in a total population of 1000. The deterministic criterion to derive this function. We can see that the function is zero between approximately .47 and .53.

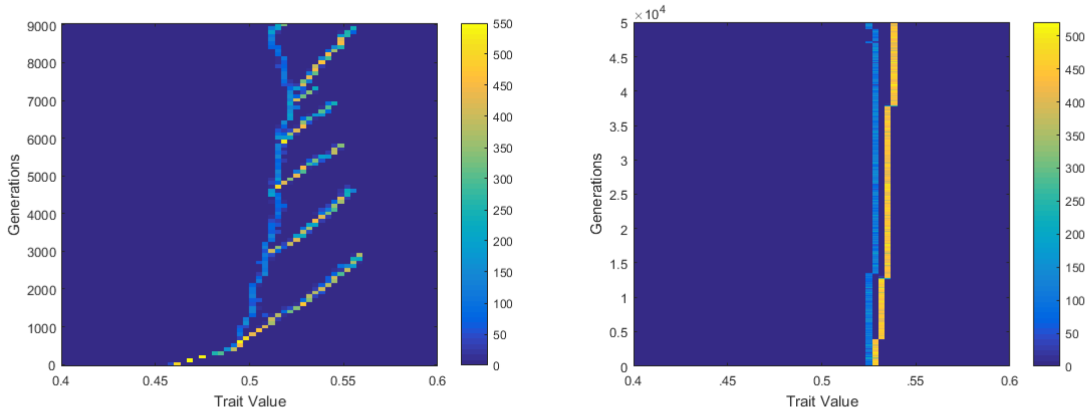


Figure C.5: Heat maps of the behavior of the simulation when there are 450 cheaters. Parameters are $\delta = .1$, $\gamma = .5$, $h = .5$, $g = .2$, $m = 6$. Left: Mutation rate is .005 with a standard deviation of .001 from the parent for the mutant offspring, with an initial population of non-cheaters with trait .46. The deterministic criterion was used in these simulations. Right: Mutation rate is .0005 with a standard deviation of .0001, with an initial population of non-cheaters at .49.

When the mutation rate is not sufficiently small, we do not see this behavior. For example, when the mutation rate is .005, there is not convergence to a single value, as shown in figure C.5 (left) although the ODE predicts convergence to approximately .47. Reducing the chance of mutation to .0005 is not enough, as evidenced by figure C.5 (right). We see that the traits pass .47 and enter the region where the ODE predicts no movement. The traits then stabilize into two coexisting branches.

Reducing the chance of mutation to .00005, leads to the behavior consistent with the ODE prediction, see figure C.6. This shows that the unusual behavior seen in figure C.5 is due to mutations occurring too often, with too large of a standard deviation of the mutant trait values. This emphasizes that our ODE prediction is accurate for sufficiently low mutation rates.

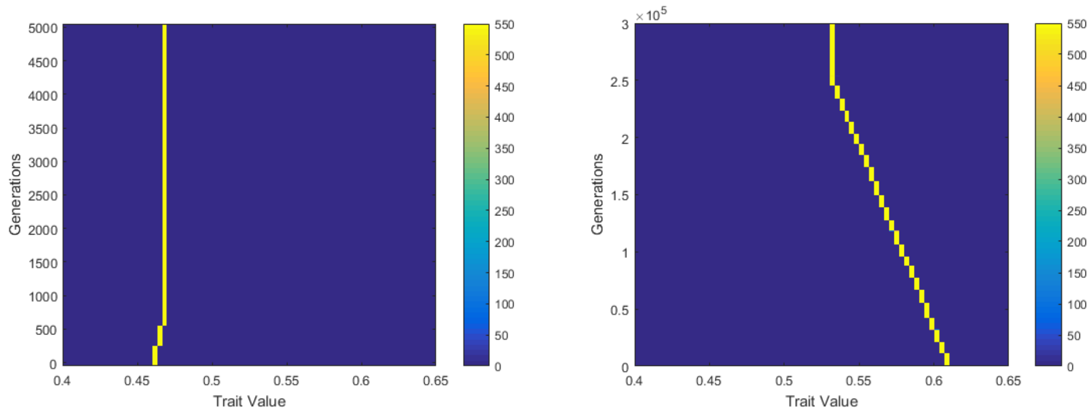


Figure C.6: Heat map of the behavior of the simulation when there are 450 cheaters and the chance of mutation is .00005 and mutations occur with a standard deviation of .0001 away from the parent. The deterministic criterion was used in these simulations. Parameters are $\delta = .1$, $\gamma = .5$, $h = .5$, $g = .2$, $m = 6$, $l = 6$, with an initial population of non-cheaters at (left) .46 and (right) .61.

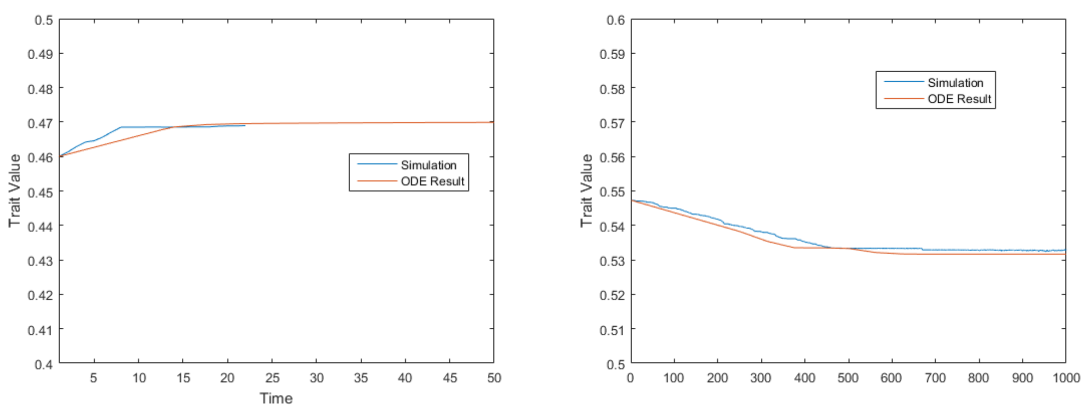


Figure C.7: Mean trait of the non-cheater population in the simulation compared with the ODE created in the case where there are 450 cheaters in a population of 1000, the chance of mutation is .00005, parameters are $\delta = .1$, $\gamma = .5$, $h = .5$, $g = .2$, $m = 6$, $l = 6$, with an initial population of non-cheaters at .46 (left) and .5472 (right). The deterministic criterion was used in the simulation and in the calculation of the the ODE.