Université de Montréal

Evolution of cooperation in evolutionary games with the opting-out strategy and under random environmental noise

par

Cong Li

Département de mathématiques et de statistique
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures et postdoctorales
en vue de l'obtention du grade de
Philosophiae Doctor (Ph.D.)
en mathématiques
Orientation mathématiques appliquées

juillet 2020

Université de Montréal

Faculté des études supérieures et postdoctorales

Cette thèse intitulée

Evolution of cooperation in evolutionary games with the opting-out strategy and under random environmental noise

Présentée par

Cong Li

A été évaluée par un jury composé des personnes suivantes

Jacques Bélair Président-rapporteur	**Sabin Lessard** Directeur de recherche
Christian Hilbe Examinateur externe	**Morgan Craig** Membre du jury
Miklós Csűrös Rep. du doyen de la FESP	

Summary

In this thesis, we study the effects of a stochastic environment and the use of an opting-out strategy on the evolution of cooperation in evolutionary games. The thesis contains 8 articles, among which 6 are already published in peer-reviewed journals. Apart from the introduction, the thesis is divided into two parts, Part 1 made with 5 articles and Part 2 with 3 articles.

Part 1 studies randomized payoffs in evolutionary games. Article 1 introduces stability concepts for 2×2 matrix games in infinite populations undergoing discrete, non-overlapping generations in a stochastic environment and gives conditions for an equilibrium, either on the boundary or in the interior of the simplex of all strategy frequencies, to be stochastically locally stable or unstable. Article 2 extends the results of Article 1 to the case where fitness is an exponential function of expected payoff in random pairwise interactions and shows that, unexpectedly, environmental random noise can break a periodic cycle and promote stability of an interior equilibrium. Article 3 discusses the effects of weak selection. While stability conditions in a random environment return to conditions in the deterministic case as selection intensity diminishes, random fluctuations in payoffs can accelerate the speed of convergence toward a stable equilibrium under weaker selection. Article 4 applies stochastic evolutionary stability theory to a randomized Prisoner's dilemma game and shows that increasing the variance in payoffs for defection is conducive to the evolution of cooperation. Article 5 studies randomized matrix games in finite populations and gives conditions for selection to favor the evolution of cooperation in the context of a randomized Prisoner's dilemma.

Part 2 considers a repeated Prisoner's dilemma game with an opting-out behavior adopted by every player in pairwise interactions. Article 6 studies the evolutionary dynamics of cooperation and defection in this context and shows possible long-term coexistence, assuming an infinite population and fast (actually, instantaneous) equilibrium in the pair frequencies. Article 7 reports experimental results with 264 university students using the opting-out strategy that support the theoretical prediction of a long-term coexistence of cooperation and defection. Article 8 extends the analysis of the model with the opting-out strategy to the case of a finite population and provides a rigorous proof of the two-time scales for the frequencies of cooperation and defection on one hand and the frequencies of strategy pairs on the other.

Keywords: Evolution of Cooperation, Randomized payoffs, Replicator dynamics, Diffusion approximation, Prisoner's Dilemma, Randomized Prisoner's Dilemma, Opting-out strategy, Fixation probability, Two-time scales, Stochastically locally stable, Stochastically locally unstable, Stochastically evolutionarily stable, Stochastically convergence stable, Weak selection, Long-term coexistence.

Sommaire

Dans cette thèse, nous étudions les effets d'un environnement stochastique et de l'utilisation d'une stratégie d'opting-out sur l'évolution de la coopération dans les jeux évolutionnaires. La thèse contient 8 articles, dont 6 sont déjà publiés dans des revues avec comité de lecture. Outre l'introduction, la thèse est divisée en deux parties, la partie 1 composée de 5 articles et la partie 2 de 3 articles.

La partie 1 étudie l'impact de gains randomisés dans les jeux évolutionnaires. L'article 1 introduit les concepts de stabilité pour les jeux avec matrice de paiement aléatoire 2×2 dans des populations infinies avec des générations discrètes sans chevauchement dans un environnement stochastique. On y donne les conditions pour qu'un équilibre, sur la frontière ou à l'intérieur du simplexe des fréquences des stratégies, soit stochastiquement localement stable ou instable. L'article 2 étend les résultats de l'article 1 au cas où la valeur sélective est une fonction exponentielle du gain attendu suite à des interactions aléatoires par paires et montre que, de manière inattendue, le bruit aléatoire environnemental peut rompre un cycle périodique et favoriser la stabilité d'un équilibre intérieur. L'article 3 discute des effets de la sélection faible. Alors que les conditions de stabilité dans un environnement aléatoire reviennent aux conditions du cas déterministe lorsque l'intensité de la sélection diminue, les fluctuations aléatoires des gains peuvent accélérer la vitesse de convergence vers un équilibre stable sous une sélection plus faible. L'article 4 applique la théorie de la stabilité évolutive stochastique à un jeu randomisé de dilemme du prisonnier. On y montre que l'augmentation de la variance des gains de défection est propice à l'évolution de la coopération. L'article 5 étudie les jeux matriciels randomisés dans des populations finies et donne les conditions pour que la sélection favorise l'évolution de la coopération dans le contexte du jeu randomisé de dilemme du prisonnier.

La partie 2 considère un jeu répété de dilemme du prisonnier dans le cas où un comportement d'opting-out est adopté par chaque joueur dans les interactions par paires.

L'article 6 étudie la dynamique évolutive de la coopération et de la défection dans ce contexte et montre une possible coexistence à long terme, en supposant une population infinie et un équilibre rapide (en fait, instantané) dans les fréquences des paires. L'article 7 rapporte des résultats expérimentaux avec 264 étudiants universitaires utilisant la stratégie d'opting-out qui soutiennent la prédiction théorique d'une coexistence à long terme de coopération et de défection. L'article 8 étend l'analyse du modèle avec la stratégie d'opting-out au cas d'une population finie et fournit une preuve rigoureuse des deux échelles de temps pour les fréquences de coopération et de défection d'une part et les fréquences de paires de stratégies d'autre part.

Mots clés: Évolution de la coopération, Gains aléatoires, Dynamique du réplicateur, Approximation de la diffusion, Dilemme du Prisonnier, Dilemme du Prisonnier alatoire, Stratégie d'opting-out, Probabilité de fixation, Deux échelles de temps, Stochastiquement localement stable, Stochastiquement localement instable, Stochastiquement évolutivement stable, Stochastiquement convergence stable, Sélection faible, Coexistence à long terme

Contents

List of abbreviations

NE	Nash Equilibrium.
ESS	Evolutionarily Stable Strategy.
CSS	Convergence Stable Strategy.
PD	Prisoner's Dilemma.
C	Cooperation.
D	Defection.
RPD	Randomized Prisoner's Dilemma.
SES	Stochastically Evolutionarily Stable.
SCS	Stochastically Convergence Stable.
SESS	Stochastically Evolutionarily Stable Strategy.
SSS	Stochastically Stable Set.
SLS	Stochastically Locally Stable.
SLU	Stochastically Locally Unstable.
IPD	Iterated Prisoner's Dilemma.
TFT	Tit-for-Tat.
AllD,	Always-defect.
AllC	Always-cooperate.
GTFT	generous Tit-for-Tat.
OFT	Out-for-Tat.
i.e.	id est (that is)
et al.	et alia or et alii (and others)
e.g.	exempli gratia (for example)
etc.	et cetera (and other similar things)

List of Figures

List of Tables

Acknowledgements

After several years of studies at University of Montreal, my Ph.D. thesis is now completed. During my stay in Canada, many people provided help for my living and my studies. Without this help, I would not have had such a wonderful time. Thus, here I would like to express my sincere gratitude towards them.

First and foremost, I want to express my deep gratitude to my supervisor, Professor Sabin Lessard, for the patient guidance in mathematics and the financial support. His rigorous academic style in mathematics had a deep influence on my research and studies. I thank him for allowing me to do some independent research based on my own interests while keeping my Ph.D. project on tract. He also helped me a lot to improve my English and writing skills.

Then, my faithful appreciation goes next to Professor Yi Tao and Xiu-Deng Zheng at Institute of Zoology, Chinese Academy of Sciences. They invited me several times to Beijing during my studies. Our collaboration keeps going and I enjoy a lot working with them. Thanks also to my other co-authors and friends, Boyu-Zhang, Ross Cressman, Lingling Deng, Ting Ji, Qiaoqiao He, Songjia Fan, Jieru Yu, Tianjiao Feng, Mingyang Wang, Shichang Wang, and all others.

I would like also to thank Professors Georges Zaccour, Iosif Polterovich, Jacques Desrosiers, who taught me courses at University of Montreal. Their teaching has helped me to extend my knowledge in a variety of fields. Thanks also to Anne-marie Dupuis, Lise Imbeault and many other administrative employees for their kindness toward me during my Ph.D. studies.

Finally, thanks to my dear friends and roommates, Ruizhi Liu, Jinxin Li, Jingyun Yang, Qifeng Gan and so on. I was glad to meet you all in Montreal and we spent a good time there together. Last but not least, I would like to express my special thanks to my parents, whose care and support motivate me to move on and make me want to be a better person.

Chapter 0

Introduction

0.1 Introduction

Cooperation is a fundamental aspect of all biological systems. It is widely observed from the organization of the cells to the behaviors of individuals. The cooperative behavior could also be described as an altruistic action where cooperators forgo some of their reproductive success to help others. Since evolution is based on fierce competition between individuals that reward only selfish behavior, the question of how natural selection can lead to cooperative behaviors is one the most important issues in evolutionary biology. In recent decades, a great deal of researches revealed the reason for the evolution of cooperation and five rules that can promote the evolution of cooperation have been proposed (Nowak, 2006a; Nowak and Sigmund, 2007). They are kin selection (Hamilton, 1964), direct reciprocity (Trivers, 1971; Axelrod and Hamilton, 1981; Axelrod, 1984), indirect reciprocity (Nowak and Sigmund, 1998a, 1998b; Ohtsuki and Iwasa, 2004; Nowak and Sigmund, 2005), network reciprocity (Nowak and May, 1992; Ohtsuki *et al.*, 2006; Li *et al.*, 2013; Zhang *et al.*, 2015) and group selection (Traulsen and Nowak, 2006; Lessard, 2011),

Among these works, a famous two-player game known as the Prisoner's Dilemma (PD) has been widely studied as one of the most important theoretical frameworks (Poundstone, 1992; Nowak and Highfield, 2011). In the PD game, two strategies, cooperation (C) and defection (D), are considered in the population and their payoffs are given by the entries of the payoff matrix $\begin{pmatrix} R & S \\ T & P \end{pmatrix}$ with the inequalities $T > R > P > S$ and $2R > T + S$, where R and S are the payoffs to C against C and D, respectively, while T and P are the payoffs to D against the same two strategies. Particularly, when $R + P = S + T$ with $P = 0$, the payoff matrix can be expressed in the form $\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}$ (Nowak, 2006b; Sigmund, 2010). In this case, the actions of C and D can be described separately. Here, cooperation represents altruism, which incurs a cost $c > 0$ to the player adopting it and provides a benefit $b > c$ to the opponent, while defection incurs no cost at all and does not disqualify from receiving a benefit from a cooperative opponent. Since the effects of C and D on payoff are additive in this model, we call it the additive PD game. On the other hand, a PD game that is repeated between the same players some number of times, fixed or random, is a called a repeated (or iterated) PD game. In particular, a one-round PD game is a PD game that is not repeated.

In the field of evolutionary game theory, stability analysis is one of the most important methods. In a population of players, a Nash equilibrium (NE) corresponds to a population state such that no player has anything to gain by changing only his own strategy. This means that no player can increase his own expected payoff by changing his strategy while the other players keep theirs unchanged. In a one-round PD game, the only NE is that every individual adopts defection (Nowak, 2006b; Sigmund, 2010). Another important concept in evolutionary game theory is that of an evolutionarily stable strategy (ESS). It is a strategy understood as a phenotype such that if every individual in the population adopts it, then any small fraction of the population adopting of mutant strategy, including a mixed strategy, cannot invade the population under the effects of natural selection (Maynard Smith and Price, 1973; Maynard Smith, 1982). In a two-player game such as a one-round PD game, a strategy represented by a frequency vector $\mathbf{x}$ is an ESS if:

(1) $E(\mathbf{y},\mathbf{x}) \leq E(\mathbf{x},\mathbf{x})$ for any strategy $\mathbf{y} \neq \mathbf{x}$; and (2) $E(\mathbf{y},\mathbf{y}) < E(\mathbf{x},\mathbf{y})$ in the case of equality in (1). Here, $E(\mathbf{y},\mathbf{x})$ represents the expected payoff to a $\mathbf{y}$-strategist against an $\mathbf{x}$-strategist. Note that condition (1) implies that an ESS in a two-player game is necessarily a NE. Moreover, in a one-round PD game, defection is simultaneously an ESS and a NE.

In an infinite population under the assumption of unlimited environmental carrying capacities, the relative growth rate of a strategy is given by its expected payoff which defines its fitness. This gives the replicator equation (Taylor and Jonker, 1978; Zeeman, 1980; Hofbauer and Sigmund, 1998) that describes the continuous-time dynamics of strategy frequencies. In the case of two strategies, the replicator equation takes the form $\dot{x} = x(1-x)(f_1(x) - f_2(x))$, where x and $1-x$ represent the frequencies of strategies 1 and 2, respectively, whose frequency-dependent fitnesses are given by $f_1(x)$ and $f_2(x)$, respectively. In generic cases, the boundary $x = 0$ ($x = 1$, respectively) is (asymptotically) stable if and only if $f_1(0) - f_2(0) < 0$ ($f_1(1) - f_2(1) > 0$, respectively). Moreover, a stable interior equilibrium $\hat{x}$ exists if and only if $f_1(\hat{x}) - f_2(\hat{x}) = 0$ for $0 < \hat{x} < 1$ and $f_1'(\hat{x}) - f_2'(\hat{x}) < 0$.

In a population of fixed finite size N, as a result of limited resources in the environment, and in the absence of mutation, the effect of random drift will bring the population state to a fixation state from any initial state. In the neutral case, that is, in the absence of selection, the probability of ultimate fixation of any strategy is given by its initial frequency. Under selection, if the fixation probability of a single mutant strategy exceeds its initial frequency $1/N$, then this strategy is said to be favored by natural selection (Nowak *et al.*, 2004). In discrete-time models, the changes in strategy frequencies are usually described by a Markov chain. In the special context of two strategies, let x be the frequency of a given strategy at time t and Δx be the change of this frequency from time t to time $t + \Delta t$ with $\Delta t = 1/N$. If, for instance, the expected value and variance of this change can be expressed as $\mathbf{E}(\Delta x) = m(x)\Delta t + o(\Delta t)$ and $\mathbf{Var}(\Delta x) = v(x)\Delta t + o(\Delta t)$ with $\mathbf{E}((\Delta x)^4) = o(\Delta t)$, then the Markov chain converges to a continuous-time continuous-space process as $N \to +\infty$. This process is actually a diffusion with drift function $m(x)$ and diffusion function $v(x)$ that satisfy Kolmogorov differential equations, from which the fixation probability can be calculated (Kimura, 1964; Ewens, 2004; Lessard, 2005). This is the approach that will be used in this thesis for finite populations.

0.2 Organization of the thesis

This rest of the thesis is organized as follows.

In section 0.3, we present some motivations. In subsection 0.3.1, we discuss the interest in considering evolutionary games in a stochastic environment and a review of the literature on games in a stochastic setting. Then we indicate how the concept of evolutionary stability in a stochastic environment can be defined. In subsection 0.3.2, we explain why the question about the coexistence of cooperation and defection in nature is so important and present some simulation results on the subject. Then we introduce the opting-out strategy as a form of direct reciprocity in mathematical models for the evolution of cooperation.

In Part 1, we study evolutionary matrix games with random payoffs. This part is made of 5 articles, among which 4 are already published and another is currently under review. In Article 1, we apply the concept of stochastic local stability to randomized 2×2 matrix games. In Article 2, we extend our results to the case where the fitness of a strategy is an exponential function of its expected payoff in random pairwise interactions. In Article 3, we explore the role of weak selection in randomized matrix games. In Articles 4 and 5, we analyse in detail randomized PD game in infinite and finite populations, respectively.

In Part 2, we study the effects that an opting-out (or Out-for-Tat) behavior in a repeated PD game may have on the evolutionary dynamics. This part contains 3 articles, 2 of which already published and the last one submitted. In Article 6, we show that cooperation and defection can coexist with the opting-out behavior adopted by every player in an infinite population assuming a faster timescale for pair frequencies than for strategy frequencies. In Article 7, we present experimental results on a repeated PD game played by 264 university students, some of whom practicing opting-out. In Article 8, we extend the model to a finite population and prove rigorously the existence of two timescales and convergence to a diffusion process in the limit of a large population.

0.3 Motivations

0.3.1 Motivation for Part 1:

In evolutionary matrix game theory, the payoffs received by the individuals are given in the form of a matrix whose entries are usually constant. However, in biological populations, individuals face uncertainty in their payoff. Such an uncertainty could possibly be introduced by probabilistic encounter rules or the use of mixed strategies (Taylor and Jonker, 1978; Eshel and Cavalli-Sforza, 1982; Hofbauer and Sigmund, 1998). On the other hand, temporal variations may also be caused by changes in the natural environment that can be periodic, such as seasonal weather fluctuations, or totally random as if occurring by accident (May, 1973; Kaplan *et al.*, 1990; Lande *et al.*, 2003).

Fluctuations in payoffs may have major effects on evolution. Of particular interest are stochastic games. These model the situation in which the environment (here, the population state) changes in response to players' choice (Shapley, 1953; Fudenberg *et al.*, 2012; Solan and Vieille, 2015; Hilbe *et al.*, 2018). In stochastic games, the action of a player has two levels of effect, the immediate effect on the payoff received and the potential effect on a future payoff influenced by the action of the player on the population state. The changes in payoffs in such games are predictable to a certain extent, being described by transition probabilities, not like random changes occurring in the surrounding environment of a natural population.

Time-dependent payoffs is another way to describe the effect of a natural environment on evolutionary games. Broom (2005) studied the case where the entries of the payoff matrix in the replicator equation in continuous time are functions of time. He compared the dynamics of the average of the population state over time to the dynamics based on the average of the payoff matrix. He showed that the time average of the population state under fluctuating payoffs can be arbitrary far from the interior Nash equilibrium for the time average payoff matrix.

Many numerical simulations of games with periodic or random payoffs can also be found. For instance, Uyttendaele *et al.* (2012) showed that stable periodic orbits, instead of stable equilibrium points, can by reached by the population state from an arbitrary initial value when the payoffs are periodic. Stollmeier and Nagler (2018) pointed out that a PD game like a Hawk-Dove game which can reach an interior stationary distribution if there is periodic noise in the payoffs. Perc (2006) investigated a PD game on a lattice network and concludes that an increase of stochastic fluctuations in payoffs may be conducive to the evolution of cooperation. Szolnoki and Perc (2019) considered periodic payoffs in four different games on a square lattice and a regular random graph. They show that periodic changes between two available games with global ordering (where cooperators survive and spread in compact clusters on the network) must be fast, while periodic changes between global and local ordering games (where cooperators and defectors are typically arranged in role-separating mixed patterns) must be slow for cooperation to thrive.

All the above theoretical and numerical works reveal that, unless small enough to be ignored, variability in payoffs may play an important role in evolutionary games. Nevertheless, this topic suffers from a lack of a theoretic mathematical framework, especially in the case of random environmental noise. The concepts of NE and ESS, originally defined in a deterministic setting, no longer work in a stochastic setting, for which extensions are missing. This was the main motivation for introducing evolutionary concepts in matrix games with random payoffs in Article 1, and studying further applications and assumptions in Articles 2-5.

0.3.2 Motivation for Part 2:

A great deal of research in evolutionary games tries to explain the evolution of cooperation. Most works related to the five rules in favor of cooperation (Nowak, 2006a; Nowak and Sigmund, 2007) focus on conditions for natural selection to lead to full cooperation in the population. In natural populations, however, it is hard to find fixation of cooperation, and at least in humans, coexistence of cooperation and defection is common (Dugatkin, 1997). The possibility of ongoing oscillations between cooperation and defection such as the alternate occurrence of war and peace periods, or cycles from C to D to TFT (see below) and neutral drift back to C in an iterated PD game are well documented (Nowak, 2006b; Sigmund, 2010). However, a more general theoretical framework is still needed to explain how long-term stable coexistence of cooperation and defection can evolve.

Under the effects of direct reciprocity, if interactions between individuals can be repeated, cooperation might prevail since clusters of cooperators helping each other would be more successful than any other (Trivers, 1971; Axelrod, 1984). In an iterated PD game, pairs of individuals play several rounds of a PD game. Cooperation could be promoted through equivalent retaliation, known as the Tit-for-Tat (TFT) strategy, starting with cooperation. This strategy goes as follows: in the first round, display C, and in the next rounds, choose the action used by the opponent in the previous round. Notice that cooperation is fixed in a population of all TFT individuals. Moreover, it has been shown by simulations that the TFT strategy does better than almost any other (Axelrod, 1984). Actually, in a game where players are restricted to either use TFT or AllD (always defect), TFT is a strict NE if the number of rounds is large enough. With random pairwise interactions in an infinite population, TFT-fixation is locally asymptotically stable, but AllD-fixation is too. Therefore, this cannot explain the evolution of cooperation when rare. In a finite population, however, the probability of ultimate fixation of a single TFT mutant exceeds its initial frequency if the frequency of TFT at the unstable interior equilibrium is below 1/3, which is the case if the number of rounds is large enough. This property has been called the one-third law of evolution (Nowak *et al.*, 2004).

In an iterated PD game, it is usually assumed that players in an interacting pair cannot decide to stop the interaction from one round to the next even if at least one of the players is not satisfied with its current payoff. As a matter of fact, the payoff received by a player in one round of a PD game is always higher against a C-opponent than against a D-opponent. Then, if the players have some control on the continuation of the game, every player should be interested in continuing an interaction with a cooperator, but prone in interrupting an interaction with a defector. As a result, CC pairs should be more robust than CD and DD pairs. This is somehow related to assortment of phenotypes that has been proposed in favor of the evolution of cooperations (Eshel and Cavalli-Sforza, 1982).

Once interrupting an interaction with a defector is available in an iterated PD game, which is called the opting-out strategy, it can be expected that this behavior will spread and finally prevail in the population. Studies on the opting-out strategy started with computer simulations, analogously to studies on the TFT strategy. For instance, Schuessler (1989) considered a population with cooperators and many types of defectors Dn that display cooperation in the first $n-1$ rounds and start to defect in round n. All strategies are "trigger strategies" so that an interaction is continued as long as both players cooperate but ended as soon as one of then defects. Here $D1$ represents the AllD strategy and $D\infty$ corresponds to cooperation with opting-out (CONCO in this article). These results show that egoistic cooperation can emerge in a population, and is more robust than expected from theoretical and sociological analyses. Hayashi (1993) presented computer simulations with 9 strategies in a "Prisoner's Dilemma network" where players are at the nodes of the network and have the option of accepting or not to interact with each other. A pair (edge) between two nodes comes into existence only by mutual acceptance of the players at the nodes. The strategies with the top 4 average payoffs in the simulation results can be considered as Out-for-Tat (OFT) strategies which are characterized by continuing to interact with a cooperating partner but seeking to interact with someone else as soon as the partner defects. The OFT strategy let the partner of the last round of interaction decide to continue or not unless the partner defects. Therefore, the OFT strategy can be considered as “TFT on partner selection”. Other computer simulations of populations on a lattice were reported in Aktipis (2004). Individuals can move to neighboring patches and an interaction occurs when two individuals are in the same patch. The “Walk Away” strategy is defined as the strategy that consists in staying on that same patch if the partner cooperates, while choosing to move if the partner defects or if there is no partner. Various strategies together with the “Walk Away” strategy are initially placed on the lattice, including pure C, pure D, TFT and PAVLOV. Moreover all these four strategies are considered in two versions, a mobile version, which consists in moving when there is no partner, and a stationary version, which consists in always staying in the same patch. It is shown that the “Walk Away” strategy is always maintained at a higher frequency in the population than any other strategy. Moreover, it can resist invasion and invade a population of defectors at a lower initial frequency than any other strategy.

On the other hand, a first theoretical analysis of the iterated PD game in an infinite population with opting-out can be found in Fujiwara-Greve and Okuno-Fujiwara (2009). The focus is put on NE conditions with various strategies which defer in the decision rules to continue or end an interaction with a partner based on the record of all previous rounds. It is shown that non-linear average payoffs may create situations with a polymorphic equilibrium where all individuals have the same expected payoff. This can support long-term coexistence of cooperation and defection in the population. Moreover, Izquierdo *et al.* (2010; 2014) presented a study of 18 strategies, which are the combinations of the player’s initial action (C or D), the response to a cooperative action (C, D or L, where L stands for leaving), and the response to a defective action (C, D or L). Through numerical simulations of an iterated PD game in an infinite population and analytical approximative mean-field methods, they find that the cornerstone of long-term stability is the coexistence of C-C-L and D-D-X where $X = D$ or L. Here C-C-L corresponds to the opting-out strategy starting with C, while D-D-D corresponds to the AllD strategy and D-D-L the opting-out strategy starting with D, whilst other strategies, such as C-C-D, which corresponds to TFT, remain present only in small population frequencies.

The above studies show the possibility of long-term coexistence of cooperation and defection in the context of an iterated PD game with the option for players to stop an interaction if this is in their best interest. However, exact conditions for this to happen had to be studied in a rigorous theoretical framework. Article 6 is a first attempt in this direction under the assumption that pair frequencies change so much faster than C and D frequencies that they are given by their equilibrium values given C and D frequencies. Article 7 shows that results of experiments conducted with university students playing an iterated PD game with the possibility of opting-out are in agreement, at least in part, with the theoretical predictions found in Article 6. Article 7 is a more rigorous mathematical treatment of the two-timescale argument that addresses the validity of a diffusion approximation for the frequency of C in an iterated PD game with opting-out in the limit of a large population.

Notice that the opting-out strategy is named differently in many papers. For instance, Schuessler (1989) calls it CONCO for conditional cooperator; Hayashi (1993) uses OFT for Out-for-Tat, in analogy with TFT for Tit-for-Tat; it is called the Walk Away strategy in Aktipis (2004) and referred to as “win stay, lose move”, by analogy to “win stay, lose shift” (Nowak and Sigmund 1993) of a PAVLOV strategy (named so in memory of the Russian psychologist “Ivan Petrovich Pavlov” who developed the concept of conditioned reflex) which consists in changing action in the next round if the opponent defects.

0.4 Main results

The thesis is made up of 8 articles. In this section, we give brief summaries of the new ideas and results contained in these articles. The full articles follow in Part 1 and Part 2. The tables 1,2 below list the assumptions that are made and the conditions we studied in each article.

Article	Population Size	Game	Model	Fitness
1	Infinite	randomized matrix game	Discrete non-overlapping	π
2	Infinite	randomized matrix game	Discrete non-overlapping	$\mathrm{Exp}(\pi)$
3	Infinite	randomized matrix game	Discrete non-overlapping	$(1-w)+w\pi$
4	Infinite	RPD	Discrete non-overlapping	π
5	Finite	randomized matrix game, RPD, iterated RPD, additive expected payoffs, additive payoffs	Wright-Fisher, Diffusion Approximation	$1+w\pi$
6, 7	Infinite, Finite	iterated PD, additive payoffs	Replicator Dynamics, Diffusion Approximation	π
8	Finite	iterated PD, additive payoffs	Moran, Diffusion Approximation	$1+w\pi$

Table 1: *Assumptions made in the articles.* PD stands for Prisoner's Dilemma, RPD for Randomized PD, and π for payoff. The parameter w represents an intensity of selection.

Article	Conditions	Equilibrium
1	SLS, SLU, SES, SCS	Boundary, Interior
2	SLS, SLU	Boundary, Interior
3	SLS, SLU, SES, Rate of convergence	Boundary, Interior
4	SLS, SLU, SES	Boundary
5	Favored by selection	Fixation
6, 7	NE	Boundary, Interior
8	Favored by selection	Fixation

Table 2: *Conditions studied in the articles.* SLS stands for stochastic local stability, SLU for stochastic local unstability, SES for stochastic evolutionary stability, SCS for stochastic convergence stability, NE for Nash equilibrium.

0.4.1 In part 1

Part 1 of the thesis contains 5 articles on matrix games in a stochastic environment. Articles 1, 2, 3 and 5 are already published, Article 4 is currently under review.

Article 1

This article introduces the concept of stochastic stability in randomized matrix games where the entries of the payoff matrix are random variables. We give conditions for a boundary or interior equilibrium to be stochastically locally stable (SLS) or stochastically locally unstable (SLU). Moreover, we extend the concepts of evolutionary stability and continuous stability originally defined in a deterministic setting to corresponding concepts in a stochastic setting.

Consider a two-strategy game in an infinite population with discrete, non-overlapping generations. The payoffs in pairwise interactions are given by the payoff matrix

$$\mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix} = \begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} . \tag{0.4.1}$$

These payoffs are assumed to be positive random variables that are uniformly bounded below and above by some positive constants. The probability distributions of $a_{ij}(t)$ for $i, j = 1, 2$ do not depend on $t \geq 0$. Moreover, a_s, b_s, c_s, d_s, a_t, b_t, c_t and d_t are all assumed to be independent for $s \neq t$. Let x_t be the frequency of strategy 1 at time $t \geq 0$. Assuming random pairwise interactions and expected payoff as fitness, the frequency of strategy 1 at time step $t + 1$ is given by the recurrence equation

$$x_{t+1} = \frac{x_t^2 a_t + x_t(1 - x_t)b_t}{x_t^2 a_t + x_t(1 - x_t)(b_t + c_t) + (1 - x_t)^2 d_t}. \tag{0.4.2}$$

We are interested in the equilibrium states of this equation and their stability properties.

According to Karlin and Liberman (1974, 1975), we introduce the following definitions of local stability or instability in a stochastic setting.

Definition 0.1 *A constant equilibrium $\hat{x}$ is said to be* stochastically locally stable *(SLS) if for any $\epsilon > 0$ there exists $\delta_0 > 0$ such that*

$$\mathbb{P}\left(x_t \to \hat{x}\right) \geq 1 - \epsilon \quad \textit{as soon as} \quad |x_0 - \hat{x}| < \delta_0 \ . \tag{0.4.3}$$

On the other hand,

Definition 0.2 *A constant equilibrium $\hat{x}$ can be said to be* stochastically locally unstable *(SLU) if*

$$\mathbb{P}\left(x_t \to \hat{x}\right) = 0 \quad \textit{as soon as} \quad |x_0 - \hat{x}| > 0 \ . \tag{0.4.4}$$

Here, $\hat{x}$ SLS means that x_t tends to $\hat{x}$ with probability arbitrarily close to 1 if the initial state x_0 is sufficiently near $\hat{x}$. On the other hand, $\hat{x}$ SLU means that with probability 1 (or almost sure), x_t does not reach $\hat{x}$ if the initial state x_0 differ from $\hat{x}$.

Applying these definitions to the two-strategy randomized game, we get the following theorem.

Theorem 0.1 *The fixation state* $\hat{x} = 0$ *of the recurrence equation (0.4.2) with the payoff matrix (0.4.1) is SLS if*

$$\mathbb{E}\left(\log\left(\frac{b_t}{d_t}\right)\right) < 0 \ , \tag{0.4.5}$$

and SLU if

$$\mathbb{E}\left(\log\left(\frac{b_t}{d_t}\right)\right) > 0 \ . \tag{0.4.6}$$

By symmetry, the fixation state $\hat{x} = 1$ *of the recurrence equation (0.4.2) with the payoff matrix (0.4.1) is SLS if*

$$\mathbb{E}\left(\log\left(\frac{c_t}{a_t}\right)\right) < 0 \ , \tag{0.4.7}$$

and SLU if

$$\mathbb{E}\left(\log\left(\frac{c_t}{a_t}\right)\right) > 0 \ . \tag{0.4.8}$$

Now, let $\bar{X}$, $\bar{Y}$, σ_X^2, σ_Y^2 and $\sigma_{X,Y}$ denote the means, variances and covariance for two random variables X and Y. When random payoffs are close enough to their means, the condition for $\hat{x} = 1$ to be SLS reduces to

$$\log\left(\frac{\bar{c}}{\bar{a}}\right) < \frac{1}{2}\left(\frac{\sigma_c^2}{\bar{c}^2} - \frac{\sigma_a^2}{\bar{a}^2}\right) . \tag{0.4.9}$$

Figure 1.1 in Article 1 gives an application of this condition to the repeated PD game. This reveals that, as the variance of the number of rounds increases, the condition for TFT-fixation to be SLS against AllD becomes more stringent.

Moreover, in the degenerate case where $b_t = d_t$ for all $t \geq 0$, it can be shown that the fixation state $\hat{x} = 0$ is SLS if

$$\mathbb{E}\left(\frac{c_t}{d_t} - \frac{a_t}{d_t}\right) = \mathbb{E}\left(\frac{c_t}{d_t}\right) - \mathbb{E}\left(\frac{a_t}{d_t}\right) \ > \ 0 \ , \tag{0.4.10}$$

and SLU if the inequality is reversed.

On the other hand, defining $u_t = x_t/(1 - x_t)$, an interior constant equilibrium $\hat{x}$ of (0.4.2) with $0 < \hat{x} < 1$ is possible only if

$$\hat{u}(a_t - c_t) = d_t - b_t. \tag{0.4.11}$$

This implies a payoff matrix in the form

$$\begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} = \begin{pmatrix} c_t + z_t & b_t \\ c_t & b_t + \hat{u}z_t \end{pmatrix} \ , \tag{0.4.12}$$

where $z_t = a_t - c_t$. Conditions for stochastic local stability or instability of the equilibrium are then given in the following theorem.

Theorem 0.2 *A constant equilibrium* $\hat{x} = \hat{u}/(1 + \hat{u})$ *of the recurrence equation (0.4.2) with* $\hat{u} > 0$ *and the payoff matrix (0.4.12) is SLS if*

$$\mathbb{E}\left(\log\left(\frac{\hat{u}a_t + d_t}{\hat{u}c_t + d_t}\right)\right) < 0 \ , \tag{0.4.13}$$

and SLU if

$$\mathbb{E}\left(\log\left(\frac{\hat{u}a_t + d_t}{\hat{u}c_t + d_t}\right)\right) > 0 \ . \tag{0.4.14}$$

In randomized matrix games, it is worth emphasizing that it is possible for a constant interior equilibrium and both fixation states to be simultaneously SLS. This is the case, for instance, with the payoff matrix

$$\begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} = \begin{pmatrix} 1 & 1 + \hat{u}\eta_t \\ 1 + \eta_t & 1 \end{pmatrix} \ , \tag{0.4.15}$$

as illustrated in Figure 1.2 in Article 1.

Moreover, the above definitions allow us to extend the concept of an evolutionarily stable strategy (ESS; Maynard Smith and Price, 1973) and the concept of a convergence stable strategy (CSS; Eshel and Motro, 1981; Eshel, 1983, 1996; Christiansen, 1991), both originally defined for evolutionary games in a constant environment, to similar concepts for evolutionary games in a stochastic environment.

Definition 0.3 *A* stochastically evolutionarily stable *(SES) strategy is a strategy such that, if all the members of the population adopt it, then the probability for at least any slightly perturbed strategy to invade the population under the influence of natural selection is arbitrarily low. More specifically, a strategy represented by a frequency vector* $\hat{\mathbf{x}}$ *is SES if* $\hat{\mathbf{x}}$*-fixation is SLS against any other strategy* $\mathbf{x} \neq \hat{\mathbf{x}}$ *at least close enough.*

Definition 0.4 *A strategy represented by a two-dimensional frequency vector given by* $\hat{\mathbf{x}} = (\hat{x}, 1 - \hat{x})$ *is said to be* stochastically convergence stable *(SCS) if the fixation state of any nearby strategy* $\tilde{\mathbf{x}} = (\tilde{x}, 1 - \tilde{x})$ *is SLU against a strategy* $\mathbf{x} = (x, 1 - x)$ *if and only if* $\mathbf{x}$ *is in the direction of* $\hat{\mathbf{x}}$ *with respect to* $\tilde{\mathbf{x}}$*. This means that* $\operatorname{sgn}(x - \tilde{x}) = \operatorname{sgn}(\hat{x} - \tilde{x})$*.*

Conditions for a boundary or interior equilibrium to be SES or SCS can be obtained.

Proposition 0.1 *Consider the positive randomized game with payoff matrix (0.4.1) whose entries are independent random variables with means* $\bar{a}_{ij}$ *and variances* σ_{ij}^2 *for* $i, j = 1, 2$*, respectively, while all higher-order centered moments are functions* $o(\sigma^2)$ *where* $\sigma^2 = \max\{\sigma_{ij}^2 : i, j = 1, 2\}$*. With the assumption that* σ^2 *is small enough and under generic conditions, the pure strategy* $\hat{\mathbf{x}} = (0, 1)$ *is SES against any nearby mixed strategy* $\mathbf{x} = (x, 1 - x)$ *if and only if*

$$\sigma_{22}^2 < \left(\frac{\bar{a}_{22}}{\bar{a}_{12}} - 1 \right) \bar{a}_{22}^2 \ . \tag{0.4.16}$$

This is also the necessary and sufficient condition for $\hat{\mathbf{x}} = (0, 1)$ *to be SCS. By symmetry, the pure strategy* $\hat{\mathbf{x}} = (1, 0)$ *is SES against any nearby mixed strategy* $\mathbf{x} = (x, 1 - x)$ *and*

SCS in generic cases under the condition that σ^2 is small enough if and only if

$$\sigma_{11}^2 < \left(\frac{\bar{a}_{11}}{\bar{a}_{21}} - 1\right)\bar{a}_{11}^2 \ . \tag{0.4.17}$$

Proposition 0.2 *Consider the positive randomized game with payoff matrix (0.4.1) where $a_{11}(t) = a_{21}(t) + z(t)$ and $a_{22}(t) = a_{12}(t) + \hat{u}z(t)$ for some constant $\hat{u} = \hat{x}/(1-\hat{x}) > 0$, with $a_{21}(t)$, $a_{12}(t)$ and $z(t)$ being independent random variables of means and variances given by $\bar{a}_{21}$, $\bar{a}_{12}$, $\bar{z}$ and σ_{21}^2, σ_{12}^2, σ_z^2, respectively, while all higher-order centered moments are functions $o(\sigma^2)$, where $\sigma^2 = \max\{\sigma_{21}^2, \sigma_{12}^2, \sigma_z^2\}$. With the assumption that σ^2 is small enough and under generic conditions, the constant mixed strategy $\hat{\mathbf{x}} = (\hat{x}, 1-\hat{x})$ is SES against any mixed strategy $\mathbf{x} = (x, 1-x)$ if and only if*

$$\bar{z}\bar{d}^2 + \bar{z}\sigma_{21}^2\hat{x}^2 + \bar{z}\sigma_{12}^2(1-\hat{x})^2 - \bar{d}\sigma_z^2\hat{x} < 0, \tag{0.4.18}$$

where

$$\bar{d} = \bar{z}\hat{x} + \bar{a}_{21}\hat{x} + \bar{a}_{12}(1-\hat{x}) > 0. \tag{0.4.19}$$

On the other hand, it is SCS under the same assumption and conditions if and only if

$$\bar{z}\bar{d}^2 + \bar{z}\sigma_{21}^2\hat{x}^2 + \bar{z}\sigma_{12}^2(1-\hat{x})^2 + \bar{z}\hat{x}^2\sigma_z^2 - \bar{d}\sigma_z^2\hat{x} < 0 \ . \tag{0.4.20}$$

These results show that an interior equilibrium that is SES and SCS can exist only with special relationships between the means, variances and covariances of the payoffs. An example is given in Figure 1.3 in Article 1. By varying the means of the payoffs but keeping fixed their variances and covariances, it is shown the situation may change from boundaries that are SES and SCS to an interior equilibrium that is SES, and finally SCS.

Article 2

In this article on evolutionary matrix games in a stochastic environment, we study the case where the fitness of a phenotype is an exponential function of the expected payoff. We show that stochastic local stability of a constant interior equilibrium can be promoted by random environmental noise even if the system may display a complicated nonlinear dynamics.

We consider a randomized game with payoff matrix (0.4.1) in an infinite population undergoing discrete, non-overlapping generations. Instead of using the expected payoff in random pairwise interactions as fitness, we use an exponential function, that is, $f_i = e^{\pi_i}$ where π_i stands for the expected payoff of strategy i for $i = 1, 2$. This is approached by the linear function $1 + \pi_i$ in the case of small payoffs, which corresponds to weak selection and leads to the replicator equation in continuous time. In the general case, the recurrence equation of x_t becomes

$$x_{t+1} = \frac{x_t e^{\pi_{1,t}}}{x_t e^{\pi_{1,t}} + (1 - x_t) e^{\pi_{2,t}}} \ , \tag{0.4.21}$$

where $\pi_{1,t} = x_t a_t + (1 - x_t) b_t$ and $\pi_{2,t} = x_t c_t + (1 - x_t) d_t$.

Extending the analysis in Article 1, conditions for a boundary equilibrium to be SLS or SLU can be found.

Theorem 0.3 *The fixation state $\hat{x} = 0$ of the recurrence equation (0.4.21) with the payoff matrix (0.4.1) is SLS if $\bar{b} - \bar{d} < 0$ and SLU if $\bar{b} - \bar{d} > 0$. By symmetry, the fixation state $\hat{x} = 1$ is SLS if $\bar{c} - \bar{a} < 0$, and SLU if $\bar{c} - \bar{a} > 0$.*

We notice that the condition for a fixation state to be SLS with a random payoff matrix is the same as the condition for the corresponding strategy to be a strict NE with the corresponding mean payoff matrix. Therefore, with exponential fitness functions, the variability in payoffs does not change the stability status of boundary equilibria. As we will see later on in Article 3, this is also consistent with results in the case of weak selection with linear fitness functions.

Next, we look at any interior equilibrium. In the absence of random environmental noise, that is, $\sigma_{ij} = 0$ for $i, j = 1, 2$ so that the payoff matrix is constant and given by its mean, an equilibrium $0 < \hat{x} < 1$ exists if $\mathrm{sgn}(\bar{b} - \bar{d}) = \mathrm{sgn}(\bar{c} - \bar{a})$, in which case $\hat{x} = (\bar{b} - \bar{d})/\gamma$, where $\gamma = \bar{b} - \bar{d} + \bar{c} - \bar{a}$. Moreover, $\hat{x}$ is globally asymptotically stable if

$0 < \gamma < 2/(\hat{x}(1-\hat{x}))$, while period doubling bifurcation and chaos occur if $\gamma > 2/(\hat{x}(1-\hat{x}))$ (Tao *et al.*, 1997). In the presence of random environmental noise, which means $\sigma_{ij} > 0$ for some i and j, a constant interior equilibrium $\hat{x}$ is possible only if there exists $\hat{u} > 0$ such that $\hat{u}(\bar{a} - \bar{c}) = \bar{d} - \bar{b}$, in which case $\hat{x} = \hat{u}/(1 + \hat{u})$. This gives the payoff matrix in the form (0.4.12). Then we have the following result.

Proposition 0.3 *With the payoff matrix (0.4.12) in the case σ_z^2 small enough and under generic conditions, a constant interior equilibrium $\hat{x} = \hat{u}/(1+\hat{u})$ of the recurrence equation (0.4.21) with $\hat{u} > 0$ is SLS if*

$$\sigma_z^2 > \left(\frac{1 + \hat{x}\bar{z}}{\hat{x}}\right) \log(1 + \hat{x}\bar{z})^2 \ , \qquad (0.4.22)$$

and SLU if the inequality is reversed.

This result shows not only that the two boundary equilibria ($\hat{x} = 0$ and $\hat{x} = 1$) and the constant interior equilibrium ($\hat{x} = \hat{u}/(1+\hat{u})$) can be simultaneously SLS, but also that an increase in the variance of the environmental noise (σ_z^2) will promote stochastic local stability of the constant interior equilibrium. This is rather counterintuitive.

Three examples are given in Article 2 and illustrated in Figures 2.1, 2.2 and 2.3. Example 1 is for the payoff matrix (0.4.12) with $\hat{u} = 1$ and shows three SLS equilibria, two on the boundary and one in the interior, in the case where $\bar{z} > 0$ and $\sigma_z^2 > (2 + \bar{z})^2 \log(1 + \bar{z}/2)^2$. Example 2 is for the same payoff matrix and shows only one stable periodic two-cycle in the case where $\bar{z} > 4$ and $\sigma_z^2 = 0$. Moreover, an increase of σ_z^2, promotes stochastic local stability of $\hat{x} = 1/2$, although a period-doubling scheme is not completely destroyed if σ_z^2 is not too large. Example 3 takes $\hat{u} = 1/2$ in the payoff matrix (0.4.12), so that $\hat{x} = 1/3$ is a constant interior equilibrium. If $\sigma_z^2 = 0$, an increase of $|\bar{z}| > 6$ can lead to period-doubling bifurcation and chaos (Tao *et al.*, 1997). However, in the case $\sigma_z^2 > 0$, for instance when $\bar{z} = -9$, it is possible to get a new peak of the probability distribution at $\hat{x} = 1/3$ instead of a stable periodic four-cycle as occurs in the case $\sigma_z = 0$.

Although nonlinear biological systems can result in the emergence of complex dynamics such as chaos (May, 1976), these are hard to find in natural populations. Our results suggest that the complexity of biological system could be reduced significantly by randomness in the surrounding environment.

Article 3

This article studies randomized matrix games in the case of weak selection. It addresses the effect of selection intensity, which changes only the speed of evolution in the deterministic case, on the evolutionary dynamics in a stochastic environment.

We consider an evolutionary matrix game in an infinite population with discrete, non-overlapping generations. With the payoff matrix in (0.4.1), the fitness of strategy i is given by in the form $f_i = (1-w) + w\pi_i$, where π_i represents the expected payoff of strategy i in random pairwise interactions, for $i = 1, 2$. The parameter w is the intensity of selection. It represents the proportion of fitness that is driven by the payoffs. Weak selection means that w is significantly small compared to 1, that is, $w \ll 1$.

With x_t denoting the frequency of strategy 1 at time t in number of generations, the recurrence equation from time t to time $t+1$ is given by

$$x_{t+1} = \frac{x_t((1-w) + w\pi_{1,t})}{x_t((1-w) + w\pi_{1,t}) + (1-x_t)((1-w) + w\pi_{2,t})} , \tag{0.4.23}$$

where $\pi_{1,t} = x_t a_t + (1-x_t)b_t$ and $\pi_{2,t} = x_t c_t + (1-x_t)d_t$. This can be viewed as a Wright-Fisher model in the limit of a large population size (Hofbauer and Sigmund, 1998), but with fitness differences of order larger than the inverse of the population size and subject to stochastic fluctuations.

A stochastic local stability analysis of the boundary equilibria yields the following result.

Proposition 0.4 *Assuming the evolutionary dynamics given by the recurrence equation (0.4.23) with the random payoff matrix (0.4.1) and under weak selection, that is,* $w \ll 1$, *the fixation state* $\hat{x} = 0$ *is SLS if* $\bar{b} - \bar{d} < 0$*, and SLU if* $\bar{b} - \bar{d} > 0$*. By symmetry, the fixation state* $\hat{x} = 1$ *is SLS if* $\bar{c} - \bar{a} < 0$*, and SLU if* $\bar{c} - \bar{a} > 0$

Moreover, if $\hat{u}(a_t - c_t) = d_t - b_t$ holds for all $t \geq 0$, where $\hat{u} = \hat{x}/(1 - \hat{x})$ is a positive constant, which means that the payoff matrix is in the form (0.4.12), we have the following result.

Proposition 0.5 *Assuming the evolutionary dynamics given by the recurrence equation (0.4.23) with the random payoff matrix (0.4.12) and under weak selection, that is,* $w \ll 1$*, the constant interior equilibrium* $\hat{x} = \hat{u}/(1 + \hat{u})$ *of the recurrence equation (0.4.23) with* $\hat{u} > 0$ *is SLS if* $\bar{c} - \bar{a} > 0$*, and SLU if* $\bar{c} - \bar{a} < 0$

These results reveal that the stability properties of an equilibrium in matrix games under weak selection in a stochastic environment and in a constant environment are similar. If an equilibrium is SLS in the randomized game under weak selection, it is a strict NE in the deterministic game with the corresponding mean payoff matrix. Moreover, there is only one exception in the reverse statement for an interior equilibrium, when $\hat{u}(a_t - c_t) = d_t - b_t$ does not hold for all $t \geq 0$. In this case, the population state in the stochastic dynamics under weak selection may tend to stay close to an interior equilibrium in the deterministic dynamics but without converging to it. Therefore, even if selection is weak, whether or not a constant interior equilibrium exists and is SLS in the stochastic dynamics cannot in general be determined only by the means of the random payoffs.

Figures 3.1 and 3.2 in Article 3 give results of numerical simulations for stochastic local stability of equilibria on the boundary and in the interior, respectively. As the intensity of selection w decreases, the evolutionary dynamics in a randomized matrix game approaches the dynamics in the deterministic game with the mean matrix payoff.

Although, variability in payoffs does not seem to affect the stability properties of equilibria in matrix games as the intensity of selection diminishes, it may have an effect on the convergence rate. As a matter of fact, when $\hat{x} = 0$ is SLS, the rate of convergence to 0 is approximated by

$$\mathbf{E}\left(\log\left[\frac{(1-w)+wb_t}{(1-w)+wd_t}\right]\right) \approx \begin{cases} w(\bar{b}-\bar{d}) + \frac{w^2\left(\sigma_d^2-\sigma_b^2\right)}{2} & \text{if } \sigma_b^2 \neq \sigma_d^2 \ , \\ w(\bar{b}-\bar{d}) + w^3\sigma^2\left(\bar{b}-\bar{d}\right) & \text{if } \sigma_b^2 = \sigma_d^2 = \sigma^2 \ , \end{cases} \tag{0.4.24}$$

while $w(\bar{b}-\bar{d}) < 0$ approximates the rate of convergence in the deterministic mean-field dynamics. Therefore, the rate of convergence in the stochastic dynamics is faster (or slower) than the rate of convergence in the deterministic mean-field approximation if $\sigma_b^2 \geq \sigma_d^2$ (or $\sigma_b^2 < \sigma_d^2$, respectively). Analogous conclusions can be drawn for the fixation state $\hat{x} = 1$.

Similarly, when the interior equilibrium $\hat{x} = \hat{u}/(1+\hat{u})$ is SLS, the rate of of convergence to the equilibrium is given by

$$\mathbf{E}\left(\log\left[\frac{\hat{u}((1-w)+wa_t)+((1-w)+wd_t)}{\hat{u}((1-w)+wc_t)+((1-w)+wd_t)}\right]\right) \approx \frac{w(\bar{a}-\bar{c})}{(1+\hat{u})} \tag{0.4.25}$$

$$+ \begin{cases} \frac{\hat{u}w^2}{2(1+\hat{u})^2}\left[\hat{u}\left(\sigma_c^2-\sigma_a^2\right)+2(\sigma_{c,d}-\sigma_{a,d})\right] & \text{if } \hat{u}\left(\sigma_c^2-\sigma_a^2\right)+2(\sigma_{c,d}-\sigma_{a,d}) \neq 0 \ , \\ \frac{w^3}{(1+\hat{u})^3}\left(\hat{u}^2\sigma_a^2+\sigma_d^2+2\hat{u}\sigma_{a,d}\right)(\bar{a}-\bar{c}) & \text{if } \hat{u}\left(\sigma_c^2-\sigma_a^2\right)+2(\sigma_{c,d}-\sigma_{a,d}) = 0 \ . \end{cases}$$

This implies that the rate of convergence to the interior equilibrium $\hat{x}$ in the stochastic dynamics is faster (or slower) than the rate in the deterministic mean-field approximation if $\hat{u}\left(\sigma_c^2-\sigma_a^2\right)+2(\sigma_{c,d}-\sigma_{a,d}) \leq 0$ (or $\hat{u}\left(\sigma_c^2-\sigma_a^2\right)+2(\sigma_{c,d}-\sigma_{a,d}) > 0$, respectively).

These last results suggest that the rate of convergence or divergence of the system near an equilibrium not only depends on environmental noise, but can even be enhanced by environmental noise. Figure 3.2b gives a example where the stochastic dynamics is faster than the deterministic mean-field dynamics.

Article 4

In this paper, we apply the results in Article 1 to study the evolution of cooperation. We consider a randomized Prisoner's Dilemma (RPD) game, which assumes a random payoff matrix whose mean determines a classic PD game, and we use the concepts of SLS and SES to show that increasing the variance in payoffs for defection is conducive to the evolution of cooperation.

Let us start with the definition of a randomized Prisoner's Dilemma (RPD) game.

Definition 0.5 *An evolutionary* 2×2 *matrix game is called a* randomized Prisoner's Dilemma *(RPD) game if the two strategies are cooperation (C) and defection (D) with random payoff matrix*

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R_t & S_t \\ T_t & P_t \end{pmatrix} \end{array} \tag{0.4.26}$$

where R_t *and* S_t *are the payoffs to* C *against* C *and* D*, respectively, while* T_t *and* P_t *are the corresponding payoffs to* D *against the same two strategies, all at time* $t \geq 0$*. Moreover, these payoffs are random variables, not necessarily independent, and their mean values determine a PD game, which means that* $\bar{T} > \bar{R} > \bar{P} > \bar{S}$ *and* $2\bar{R} > \bar{T} + \bar{S}$*.*

We consider a RPD game with positive and bounded payoffs in an infinite population with discrete, non-overlapping generations. Let x_t and $1 - x_t$ denote the frequencies of C and D, respectively, at time $t \geq 0$ in number of generations. Assuming random pairwise interactions, the recurrence equation from time t to time $t + 1$ is given by

$$x_{t+1} = \frac{x_t \pi_{C,t}}{x_t \pi_{C,t} + (1 - x_t)\pi_{D,t}} \ , \tag{0.4.27}$$

where $\pi_{C,t} = x_t R_t + (1 - x_t) S_t$ and $\pi_{D,t} = x_t T_t + (1 - x_t) P_t$. Applying Theorem 0.1 for a boundary equilibrium and assuming that the variances of the payoffs are small, the following result can be shown.

Proposition 0.6 *Assuming the recurrence equation (0.4.27) for the frequency of* C *in a RPD game with payoff matrix (0.4.26) in the case of small enough variances in the payoffs,* C*-fixation is SLS if*

$$\log\left(\bar{R}/\bar{T}\right) > \left(\sigma_R^2/\bar{R}^2 - \sigma_T^2/\bar{T}^2\right)/2 \ , \tag{0.4.28}$$

and SLU if the inequality is reversed. By symmetry, D-fixation is SLS if

$$\log\left(\bar{P}/\bar{S}\right) > \left(\sigma_P^2/\bar{P}^2 - \sigma_S^2/\bar{S}^2\right)/2 \ , \tag{0.4.29}$$

and SLU if the inequality is reversed.

In particular, C-fixation is SLS if $\sigma_R^2 = 0$ and $\sigma_T^2 > -2\bar{T}^2 \log\left(\bar{R}/\bar{T}\right)$, while D-fixation is SLU if $\sigma_S^2 = 0$ and $\sigma_P^2 > 2\bar{P}^2 \log\left(\bar{P}/\bar{S}\right)$. Therefore, under the RPD game dynamics, not only C-fixation can be SLS even if $\bar{T} > \bar{R}$, but also D-fixation can be SLU even if $\bar{P} > \bar{S}$. This implies that an increase of σ_T^2 may promote C-fixation, while an increase of σ_P^2 may push the population state away from D-fixation.

Figure 4.1a,b in Article 4 presents simulation results on the RPD game obtained by varying the values of $\bar{T}$, $\bar{P}$, σ_T^2 and σ_P^2. These are in agreement with the theoretic predictions, even if the variances of the payoffs are not that small.

On the other hand, conditions for C and D to be SES can be obtained.

Proposition 0.7 *Under the assumptions of the previous proposition, the pure strategy C is SES against any nearby mixed strategy if we have $\bar{R}^2(\bar{T} - \bar{R}) > -(\bar{T}\sigma_R^2 - \bar{R}\sigma_{RT})$. By symmetry, the pure strategy D is SES against any nearby mixed strategy if we have $\bar{P}^2(\bar{S} - \bar{P}) < -(\bar{S}\sigma_P^2 - \bar{P}\sigma_{SP})$.*

We notice that, as the covariances between the payoffs diminish to 0, the conditions for C and D to be SES are in agreement with the conditions for pure strategies in randomized 2×2 matrix games with independent payoffs to be SES as given in proposition 0.1. Moreover, combining the conditions in the two propositions above, we see that there exists a threshold value of the ratio $\bar{P}/\bar{S}$, denoted by z^*, such that D-fixation is SLS if $\bar{P}/\bar{S} < z^*$, while the pure strategy D may not be SES. Conversely, D-fixation is SLU if $\bar{P}/\bar{S} > z^*$, while the pure strategy D may be SES. This is illustrated in Figure 4.1c in Article 4.

Article 5

In this article, we study a randomized 2×2 matrix game in a finite population assuming a Wright-Fisher reproduction scheme (Fisher, 1930; Wright, 1931; Cannings, 1974). We establish a diffusion approximation in the limit of a large population and apply the approximation in the case of the RPD game. This will allow us to show that increasing the variance in payoffs for defection should increase the probability of ultimate fixation of a single cooperating mutant.

We consider discrete, non-overlapping generations in a finite population of fixed size N and measure time in number of N generations. Then, $\Delta t = 1/N$ represents the time interval from one generation to the next. There are two possible pure strategies in the population, S_1 and S_2, and the payoffs at time $t \geq 0$ are given by the matrix

$$\begin{array}{cc} & \begin{array}{cc} S_1 & S_2 \end{array} \\ \begin{array}{c} S_1 \\ S_2 \end{array} & \begin{pmatrix} \eta_1(t) & \eta_2(t) \\ \eta_3(t) & \eta_4(t) \end{pmatrix} \end{array}, \tag{0.4.30}$$

where $\eta_i(t)$ are random variables with values that are always larger than -1 and probability distributions that do not depend on time $t \geq 0$, for $i = 1, 2$.

In addition, we assume that these payoffs have expected values, variances and covariances of order given by the inverse of the population size. More precisely, they can be expressed in the form $\eta_i(t) = \mu_i N^{-1} + \xi_i(t)$, where $\mathbf{E}(\xi_i(t)) = 0$, $\mathbf{Var}(\xi_i(t)) = \sigma_i^2 N^{-1}$ and $\mathbf{Cov}(\xi_i(t), \xi_j(t)) = \sigma_{ij} N^{-1}$, while all the higher-order moments of $\xi_i(t)$ are functions $o(N^{-1})$, for $i, j = 1, \ldots, 4$ with $i \neq j$. The population-scales parameters μ_i, σ_i^2 and σ_{ij} are assumed constant.

We suppose that the expected payoffs in random pairwise interactions have additive effects on fitness understood as relative reproductive success with a baseline value equal to 1. Given that the frequency of strategy S_1 is $x(t)$ at time $t \geq 0$, and as a result of a Wright-Fisher reproduction model, $x(t + \Delta t)$ follows a binomial distribution whose parameters are the population size N and a random frequency

$$x'(t) = \frac{x(t)\pi_1(t)}{x(t)\pi_1(t) + (1 - x(t))\pi_2(t)} \ , \tag{0.4.31}$$

where $\pi_1(t) = 1 + x(t)\eta_1(t) + (1 - x(t))\eta_2(t)$ and $\pi_2(t) = 1 + x(t)\eta_3(t) + (1 - x(t))\eta_4(t)$.

Then the change in the frequency of S_1 from time t to time $t + \Delta t$, which is given by $\Delta x = x(t + \Delta t) - x(t)$, satisfies

$$\mathbf{E}(\Delta x | x(t) = x) = m(x)\Delta t + o(\Delta t), \tag{0.4.32}$$

$$\mathbf{E}((\Delta x)^2 | x(t) = x) = v(x)\Delta t + o(\Delta t), \tag{0.4.33}$$

$$\mathbf{E}((\Delta x)^4 | x(t) = x) = o(\Delta t), \tag{0.4.34}$$

where

$$\begin{aligned} m(x) = x(1-x)\Big(&\mu_2 - \mu_4 + x\big(\mu_1 - \mu_2 - \mu_3 + \mu_4\big) \\ &+ x^3(\sigma_{13} - \sigma_1^2) + x(1-x)^2(2\sigma_{34} - \sigma_{14} - \sigma_{23} + \sigma_{24} - \sigma_2^2) \\ &+ x^2(1-x)(-2\sigma_{12} + \sigma_{14} + \sigma_{23} - \sigma_{13} + \sigma_3^2) + (1-x)^3(\sigma_4^2 - \sigma_{24})\Big) \end{aligned} \tag{0.4.35}$$

and

$$\begin{aligned} v(x) = x(1-x)\Big(&1 + x^3(1-x)(\sigma_1^2 + \sigma_3^2 - 2\sigma_{13}) + x(1-x)^3(\sigma_2^2 + \sigma_4^2 - 2\sigma_{24}) \\ &+ 2x^2(1-x)^2(\sigma_{12} + \sigma_{34} - \sigma_{14} - \sigma_{23})\Big). \end{aligned} \tag{0.4.36}$$

These conditions ascertain a diffusion approximation with drift function $m(x)$ and diffusion function $v(x)$ in the limit of a large population with the population size N as unit of time (see, e.g., Kimura, 1964; Risken, 1992; Ewens, 2004). Owing to diffusion theory, the probability of ultimate fixation of strategy S_1 given an initial frequency $x(0) = p$ is given by

$$u(p) = \frac{\int_0^p \psi(y)dy}{\int_0^1 \psi(y)dy}, \tag{0.4.37}$$

where

$$\psi(y) = \exp\left(-2\int_0^y \frac{m(x)}{v(x)}dx\right) \tag{0.4.38}$$

with $m(x)$ and $v(x)$ given above.

Considering the case of a RPD game with independent payoffs for cooperation (C) or defection (D) against cooperation or defection, that is $\sigma_{ij} = 0$ for $i, j = 1, \ldots, 4$ with $i \neq j$, the population-scaled expected payoffs are given by

$$\begin{pmatrix} \mu_1 & \mu_2 \\ \mu_3 & \mu_4 \end{pmatrix} = \begin{pmatrix} R & S \\ T & P \end{pmatrix} \tag{0.4.39}$$

with $T > R > P > S$ and $2R > T + S$. Now, let $F_C = u(1/N)$ ($F_D = 1 - u(1 - 1/N)$, respectively) denote the probability of ultimate fixation of a single cooperating mutant (defecting mutant, respectively) in an all defecting population (cooperating population, respectively). We say that *the evolution of cooperation is favored by selection*, *the evolution of cooperation is disfavored (not favored) by selection* and *the evolution of cooperation is more favored by selection than the evolution of defection*, if the conditions $F_C > 1/N$, $F_D < 1/N$ and $F_C > F_D$, respectively, are satisfied. Assuming the population size N large enough, these conditions are equivalent to the conditions $\int_0^1 \psi(y)dy < 1$, $\int_0^1 \psi(y)dy > \psi(1)$ and $\psi(1) < 1$, respectively.

It can be shown that the partial derivatives of $g(x) = m(x)/v(x)$ with respect to the variances of the payoffs satisfy $\partial g(x)/\partial\sigma_3^2 > 0$ and $\partial g(x)/\partial\sigma_4^2 > 0$ for $0 < x < 1$. This leads to the following result.

Proposition 0.8 *Assuming a Wright-Fisher model for a RPD game with independent payoffs in a large but finite population, whose payoffs given in (0.4.30) have population-scaled expected values given in (0.4.39), increasing the variance of at least one payoff for defection, this is* σ_3^2 *and* σ_4^2*, increases the probability of ultimate fixation of cooperation introduced as a single mutant in an all defecting population,* F_C*, while it decreases the probability of ultimate fixation of defection introduced as a single mutant in an all cooperating population,* F_D*.*

Now, consider the situation where the population-scaled expected payoffs determine an additive PD game, that is,

$$\begin{pmatrix} \mu_1 & \mu_2 \\ \mu_3 & \mu_4 \end{pmatrix} = \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \tag{0.4.40}$$

with $b > c > 0$, where c represents a cost of cooperation and b a benefit. When $c \leq 1$, it is possible to show that $\partial g(x)/\partial \sigma_1^2 < 0$ and $\partial g(x)/\partial \sigma_2^2 < 0$ for $0 < x < 1$. Thus we get the following supplementary result.

Proposition 0.9 *Under the assumptions of the previous proposition for a RPD game but with additive expected payoffs so that the population-scaled values are in the form (0.4.40) where $c \leq 1$, diminishing the variance of at least one payoff for cooperation, that is, σ_1^2 or σ_2^2, increases the probability of ultimate fixation of cooperation introduced as a single mutant in an all defecting population, F_C, while it decreases the probability of ultimate fixation of defection introduced as a single mutant in an all cooperating population, F_D.*

In section 5.5 of Article 5, we discuss five special cases of the RPD game with additive expected payoffs to exhibit the effects of the variances on conditions under which the evolution of cooperation could be favored by selection. Table 5.1 gives the numerical threshold values for $F_C > 1/N$, $F_D < 1/N$ and $F_C > F_D$ to hold.

Next, we turn our attention to the case of a RPD game with additive payoffs that are not independent. We suppose that the cost c and benefit b in an additive deterministic PD game are replaced with random variables $c(t)$ and $b(t)$, respectively, so that the payoff matrix at time $t \geq 0$ is given by

$$\begin{pmatrix} \eta_1 & \eta_2 \\ \eta_3 & \eta_4 \end{pmatrix} = \begin{pmatrix} b(t)-c(t) & -c(t) \\ b(t) & 0 \end{pmatrix} . \tag{0.4.41}$$

Moreover, let μ_c, μ_b, σ_c^2, σ_b^2 and σ_{bc} represent the population-scaled expected values, variances and covariance of $c(t)$ and $b(t)$, respectively. Then, the conditions for $F_C > 1/N$, $F_D < 1/N$ and $F_C > F_D$ become

$$\sigma_{bc} - \sigma_c^2 > 3\mu_c, \quad \sigma_{bc} - \sigma_c^2 > 2\mu_c, \quad \sigma_{bc} - \sigma_c^2 > \frac{3}{2}\mu_c, \tag{0.4.42}$$

respectively. Since $\mu_c > 0$, these conditions can hold only if $\sigma_{bc} > \sigma_c^2$, in which case the first condition is the most stringent one and the third condition the least stringent one.

In particular, if $b(t) = rc(t)$ where $r > 0$ is a constant that represents the benefit to cost ratio, the above conditions reduce to

$$r > 1 + 3\left(\frac{\mu_c}{\sigma_c^2}\right), \quad r > 1 + 2\left(\frac{\mu_c}{\sigma_c^2}\right), \quad r > 1 + \frac{3}{2}\left(\frac{\mu_c}{\sigma_c^2}\right), \tag{0.4.43}$$

respectively.

The above results can be extended in the case of an iterated RPD game, which is done in Section 5.7 of Article 5.

0.4.2 In part 2

Part 2 of the thesis contains 3 articles on the effects of opting-out on the evolution of cooperation. Articles 6 and 7 are already published, while Article 8 is under review.

Article 6

Article 6 studies the iterated PD game with opting-out in a large population under an assumption of two time-scales. It provides a theoretical framework for explaining long-term coexistence of cooperation and defection.

We consider a PD game with additive payoffs for cooperation (C) and defection (D) in pairwise interactions given by the 2×2 matrix

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array} \cdot \tag{0.4.44}$$

Every interaction between two players can be repeated but, at the end of each round, each player can unilaterally decide whether or not to continue the interaction with the same player. Based on self-interest in the PD game, both a cooperator and a defector prefer to interact with a cooperator. Thus, the rational choice of every player is to continue an interaction if the opponent is a cooperator and to stop it if the opponent is a defector. This is called the opting-out strategy, or Out-for-Tat (OFT) strategy for short. Moreover, it is assumed that, even if two players are willing to continue their interaction, this interaction may stop at the end of each round with some fixed probability $\rho > 0$. In all cases, when an interaction between two players stops for one reason or another, the players become free to form new pairs for the next round (see Figure 7.3 for a setup of the model).

Assuming random pairing of free players at the end of each round of interaction, the pair frequencies, denoted by P_{CC}, P_{CD} and P_{DD}, have expected changes from one round to the next given by

$$\mathbf{E}\left(\Delta P_{CC}\right) = (1-\rho)P_{CC} + \frac{(2\rho P_{CC} + P_{CD})^2}{4(1-(1-\rho)P_{CC})} - P_{CC} \ , \tag{0.4.45a}$$

$$\mathbf{E}\left(\Delta P_{CD}\right) = \frac{(2\rho P_{CC} + P_{CD})(P_{CD} + 2P_{DD})}{2(1-(1-\rho)P_{CC})} - P_{CD} \ , \tag{0.4.45b}$$

$$\mathbf{E}\left(\Delta P_{DD}\right) = \frac{(P_{CD} + 2P_{DD})^2}{4(1-(1-\rho)P_{CC})} - P_{DD} \ , \tag{0.4.45c}$$

if the population is large enough and selection is neglected. On the other hand, the frequency of C is given by $x = P_{CC} + P_{CD}/2$. Keeping this frequency fixed and solving $\mathbf{E}\left(\Delta P_{CC}\right) = \mathbf{E}\left(\Delta P_{CD}\right) = \mathbf{E}\left(\Delta P_{DD}\right) = 0$, we get the equilibrium condition $P_{CD}^2 = 4\rho P_{CC}P_{DD}$. This condition assumes that the changes in C and D frequencies occur at a longer time-scale than the changes in pair frequencies. It is similar to the Hardy-Weinberg equilibrium in population genetics (Ewens, 2004).

Using the fact that $P_{CC} + P_{CD} + P_{DD} = 1$, the equilibrium condition becomes

$$P_{CD}^* = -\frac{\rho}{1-\rho} + \sqrt{\left(\frac{\rho}{1-\rho}\right)^2 + \frac{4x(1-x)\rho}{1-\rho}} \ . \tag{0.4.46}$$

This function is assumed to give the frequency of CD pairs given a frequency x for C and $1-x$ for D (see Figure 7.4a in Article 6 for an illustration). Then, the expected payoffs of C and D are given by

$$\pi_C = \frac{2x - P_{CD}^*}{2x} b - c \ , \tag{0.4.47a}$$

$$\pi_D = \frac{P_{CD}^*}{2(1-x)} b \ . \tag{0.4.47b}$$

These are used in the replicator equation for C and D in an infinite population in continuous time to yield

$$\frac{dx}{dt} = x(1-x)(b-c) - \frac{bP^*_{CD}}{2} \ . \tag{0.4.48}$$

With respect to this dynamics, the boundary $x = 0$ is locally asymptotically stable and the boundary $x = 1$ is unstable. On the other hand, if $\rho < (b-c)^2/(b+c)^2$, there are two interior equilibria, denoted by x_1^* and x_2^* with $0 < x_2^* < 1/2 < x_1^* < 1$. Their expressions are given by

$$x^*_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{bc}{(b-c)^2}\frac{\rho}{1-\rho}} \ . \tag{0.4.49}$$

Here, the interior equilibrium x_1^* (with $x_1^* > 1/2$) is locally asymptotically stable, while x_2^* (with $x_2^* < 1/2$) is unstable. If $\rho = (b-c)^2/(b+c)^2$, then $x_1^* = x_2^* = 1/2$ is an unstable equilibrium. Finally, if $\rho > (b-c)^2/(b+c)^2$, then no interior equilibrium exists and $x = 0$ is the only globally stable equilibrium. Figure 7.4b in Article 6 summarizes the phase portrait of the dynamics.

Moreover, in the iterated PD game with OFT used by all players, none of the mixed strategy $(x, 1-x)$, even $(x_1^*, 1-x_1^*)$, can be an ESS when $\rho < (b-c)^2/(b+c)^2$. However, when the population is at the interior stable equilibrium x_1^*, no individual not using OFT can successfully invade this population. This is illustrated by simulation results that show that neither AllD nor TFT can invade a mixed population where they must play against both OFT-cooperators and OFT-defectors (see Figure 6.1 in Article 6 for details).

Article 7

Article 7 reports results of an experiment conducted with 264 university students playing an iterated PD game with or without opting-out. Note that none of the participants had any background in game theory or economics, they all played anonymously via computer screens, and they were not allowed to communicate with each other.

The participants in the experiment were divided into five groups, including two control groups (C1 and C2) and three treatment groups (T1, T2 and T3). All three treatment groups used the same setting, thus we treated them as a single group, denoted by T. The payoff matrix in one round of a game between two players using either cooperation (C) or defection (D) as strategy is given in all cases by

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 4 & 1 \\ 5 & 2 \end{pmatrix} \end{array}, \tag{0.4.50}$$

which corresponds to a PD game with cost $c = 1$ and benefit $b = 3$ for cooperation, subtracted from, or added to, a baseline payoff 2 to avoid negative values. Games without any opting-out were played in the control groups C1 and C2, an iterated PD game in C1, with a fixed probability 5/6 of repetition of the game between the same players, and a one-round PD game in C2. An iterated PD game with opting-out was played in the treatment group T. In this group, at the end of each round, an interaction between two players is continued with probability 5/6 but only if both player are willing to continue. Otherwise, the interaction was interrupted. When an interaction between two players was interrupted, the players form new pairs at random with players in the same situation.

The experimental results presented in Figure 7.1 in Article 7 concern the cooperation level, defined as the frequency that C is used. Comparing the averages of this frequency for the first 10 rounds and the last 10 rounds, we notice an increase from 0.64 to 0.8 in C1, a drop from 0.39 to 0.28 in C2 and non-significant increase from 0.56 to 0.58 in T. These results are basically in agreement with the equilibrium structures in the theoretical model, where $x = 0$ is the only NE in C2, while $x = 1$ and $x = x_1^* = 0.82$ are locally asymptotically stable in C1 and T. However, the values of the equilibrium points are not exactly the same. In the theoretical analysis, we assumed that players always had a rational behavior and time was long enough for the population to evolve to an equilibrium, while in the experiment, participants may act irrationally and the number of rounds is not large. Nevertheless, the experiment provides a possible explanation for a stable coexistence of C and D.

Moreover, in the treatment group T, the proportion of players continuing the interaction with the opponent in the next round is 0.92 if the opponent displays C, whereas this proportion drops to 0.53 if the opponent displays D. More precisely, a C-player in a CC pair chooses to continue with probability 0.94, while a C-player (D-player, respectively) in a CD pair chooses to continue with probability 0.67 (0.65, respectively), and a D-player in a DD pair chooses to continue with probability 0.57 (see Figure 7.2 in Article 7). The action of a cooperator in CC pairs supports our assumption on opting-out, since then a C-player chooses to continue with probability close to 1. However, the actions of the other players are not in agreement with our assumption, which can explain the discrepancy between the observed equilibrium values in the experiment and the values predicted from the theoretical model.

There might be several possible reasons for the actual behavior of the participants in the experiment. The frequency of C-players among all the free players, who did not repeat the interaction with their partner, is much less than the frequency of C in the population. The probability to find a new partner displaying C is not that high. Moreover, we surmise that the following factors might come into play:

(1) a C-player in a CD pair expects a change of strategy of the opponent or seeks a revenge against the opponent in the next round;

(2) a D-player in CD pair fears a revenge of the opponent in the next round.

On the other hand, we observed that 85.16% of the participants in the experiment stopped their interaction with their opponent with significantly higher probability when the opponent displayed D than when the opponent displayed C. These participants can be seen as OFT-preferred strategists. On the contrary, only 3.3% of the participants stopped their interaction with significantly higher probability when the opponent displayed C than when it displays D.

Article 8

Article 8 studied the iterated PD game with opting-out in a large but finite population with Moran process (Hofbauer and Sigmund, 1998; Ewens, 2004). Scalings of the intensity of selection and time with respect to the population size, we will establish a diffusion approximation with two time scale problem, where pair frequency change faster than strategy frequency. Moreover, for an additive PD game, we give the conditions for cooperation to be favored by selection.

We consider a population consist of N pairs of interacting individuals of strategies cooperation C or defection D. And the payoffs are given by the entries of a 2×2 matrix.

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} \pi_{CC} & \pi_{CD} \\ \pi_{DC} & \pi_{DD} \end{pmatrix} \end{array}, \tag{0.4.51}$$

The change of population states from t to $t + \Delta t$, where $\Delta t = 1/(2N^2)$, contains a recombination process of pairs and a birth-death update. The process is illustrated in Figure 8.1. At the beginning of time t, suppose that all individual adopt opting-out strategy, so that all CD and DD pairs break apart, while each CC pair breaks apart with probability ρ, which is a constant. The free individuals who leave their previous interactions will reform pairs at random. After that, the frequency of CC, CD and DD in all pairs are denoted by q_{CC}, q_{CD} and q_{DD} respectively, and the expression of their expectations are given in (8.2.5). The fitness of an i-strategist when interacting with a j-strategist is given by

$$w_{ij} = 1 + s\pi_{ij} \tag{0.4.52}$$

where $s = \sigma/N$ represents the intensity of selection. Then one individual is chosen with probability proportional to the fitness to produce an offspring identical to itself and one individual is chosen at random to be replaced by the offspring.

Following the replacement of an individual by the offspring, the frequencies of CC, CD and DD among the N pairs are denoted by P'_{CC}, P'_{CD} and P'_{DD} respectively, and the frequency of C in these pairs is $P'_{CC}+P'_{CD}/2 = x'$. Let $\Delta x = x'-x$ and $\Delta P_{CC} = P'_{CC}-P_{CC}$ denote the change in frequency of C and CC respectively. Given these frequencies at time t, the conditional moments of Δx and ΔP_{CC} are approximated as

$$\mathbf{E}(\Delta x) = \frac{\sigma}{2N^2}\mathbf{E}\left((1-x)A + xB\right) + o(N^{-2}), \tag{0.4.53a}$$

$$\mathbf{E}((\Delta x)^2) = \frac{x(1-x)}{2N^2} + o(N^{-2}), \tag{0.4.53b}$$

$$\mathbf{E}((\Delta x)^4) = o(N^{-3}), \tag{0.4.53c}$$

where $A = q_{CC}\pi_{CC} + \frac{q_{CD}}{2}\pi_{CD}$ and $B = q_{DD}\pi_{DD} + \frac{q_{CD}}{2}\pi_{DC}$, and

$$\mathbf{E}(\Delta P_{CC}) = \frac{(x-P_{CC})^2 - \rho P_{CC}(1-2x+P_{CC})}{1-(1-\rho)P_{CC}} + O(N^{-1/2}), \tag{0.4.54a}$$

$$\mathbf{Var}(\Delta P_{CC}) = O(N^{-1}). \tag{0.4.54b}$$

Moreover, in an infinite population in the absence of selection, the frequency of C remains constant while the frequency of CC converges uniformly to an equilibrium value

$$P^*_{CC} = x + \frac{\rho}{2(1-\rho)} - \frac{\sqrt{\rho^2 + 4x(1-x)\rho(1-\rho)}}{2(1-\rho)}, \tag{0.4.55}$$

These conditions (0.4.53), (0.4.54) and (0.4.55) show that there are two time scales in the discrete-time Markov chain for the population state, where variable P_{CC} changes more rapidly than variable x. Substituting $P_{CC} = P^*_{CC}$ in the expression of q_{CC}, q_{CD} and q_{DD} leads to

$$\mathbf{E}(q_{CC}) = P^*_{CC} + O(N^{-1/2}), \tag{0.4.56a}$$

$$\mathbf{E}(q_{CD}) = 2x - 2P^*_{CC} + O(N^{-1/2}), \tag{0.4.56b}$$

$$\mathbf{E}(q_{DD}) = 1 - 2x + P^*_{CC} + O(N^{-1/2}). \tag{0.4.56c}$$

As $N \to \infty$, this gives the result below according to Ethier and Nagylaki 1980.

Proposition 0.10 *Consider a PD game with payoff matrix (0.4.51) for N pairs of individuals so that, as a result of opting-out from one round to the next, all pairs break apart to form new pairs at random but a random proportion of CC pairs whose mean is $1-\rho < 1$.*

Assume one birth-death event at the end of each round with the probability of giving birth proportional to an affine function of payoff with coefficient σ/N *and the probability of dying given by* $1/(2N)$*. Taking* $2N^2$ *birth-death events as unit of time and letting* $N \to \infty$*, the Markov chain of the frequency of* C *converges to a diffusion with* $v(x) = x(1-x)$ *as diffusion function and*

$$m(x) = \sigma\left(x(1-x)(\pi_{CC} - \pi_{DD}) - (x - P^*_{CC})\left((1-x)(\pi_{CC} - \pi_{CD}) + x(\pi_{DC} - \pi_{DD})\right)\right) \tag{0.4.57}$$

as drift function, where P^*_{CC} *is given by (0.4.55).*

Now, considering an additive PD game with payoff matrix (0.4.44), the drift function above can be simplified to

$$m(x) = \sigma\left[x(1-x)(b-c) - \frac{bf(x,\rho)}{2(1-\rho)}\right], \tag{0.4.58}$$

where $f(x,\rho) = \sqrt{\rho^2 + 4x(1-x)\rho(1-\rho)} - \rho$. Let F_C (F_D respectively) be the probability of ultimate fixation of C (D respectively) introduced as a single mutant in an all D (C respectively) population. Following the terminology in Article 5, selection favors the evolution of cooperation, favors the evolution of cooperation more than defection, disfavors the evolution of defection if and only if $F_C > (2N)^{-1}$, $F_C > F_D$ and $F_D < (2N)^{-1}$ respectively. These conditions are also equivalent to

$$\int_0^1 \psi(y)dy < 1, \tag{0.4.59a}$$

$$\psi(1) < 1, \tag{0.4.59b}$$

$$\int_0^1 \psi(y)dy > \psi(1), \tag{0.4.59c}$$

respectively, where $\psi(y) = \exp\left(-2\int_0^y \frac{m(x)}{v(x)}dx\right)$. With the notation $r = b/c > 1$ for the benefit-to-cost ratio and $a = (1/\rho) - 1 \geq 0$ for the expected number of times that each CC pair continues to interact, we can get the proposition from the monotonicity of $\psi(y)$

Proposition 0.11 *Consider an additive PD game with payoff matrix (0.4.44) in the framework of proposition 0.10 with* $0 < \rho < 1$*. In a large enough population, increasing the value of* $r = b/c > 1$ *or* $a = 1/\rho - 1 > 0$ *(or decreasing the value of* ρ*) increases (decreases, respectively) the probability of ultimate fixation of cooperation (defection, respectively) introduced as a single mutant in an all defecting (cooperating, respectively) population,* F_C *(*F_D*, respectively).*

Moreover, when $F_C = F_D$, it is possible to show that $F_C = F_D < (2N)^{-1}$. Combined with the proposition above, we get the following result

Proposition 0.12 *In the setting of proposition 0.11, as the value of* r *or* a *increases, the conditions (0.4.59c), (0.4.59b) and (0.4.59a) for* $F_D < (2N)^{-1}$*,* $F_C > F_D$ *and* $F_C > (2N)^{-1}$*, respectively, are satisfied in this order. In particular, when cooperation is favored by selection, it is necessarily fully favored by selection.*

Finally, in order to get explicit conditions for cooperation to be favored by selection, we use the inequalities

$$4x(1-x)(\sqrt{\rho}-\rho) \leq f(x,\rho) \leq \sqrt{4x(1-x)}(\sqrt{\rho}-\rho), \tag{0.4.60}$$

and replace $f(x,\rho)$ in $\psi(y)$ with the lower bound as its approximation. Thus we get condition for cooperation to be favored by selection in any sense, that is $r > (1+\sqrt{\rho})/(1-\sqrt{\rho})$

Figure 8.2 illustrates the surfaces determined by $f(x,\rho)$ and its upper, lower bound. It approaches to the lower bound as $\rho \to 0$ while approach to the upper bound as $\rho \to 1$. Figure 8.3 shows the difference between the exact numerical solution of $F_D = (2N)^{-1}$, $F_C = F_D$, $F_C = (2N)^{-1}$ and the approximation obtained by using the lower and upper bounds of $f(x,\rho)$.

Bibliography

[1] C.A. Aktipis, Know when to walk away: contingent movement and the evolution of cooperation. *Journal of Theoretical Biology* **231**, 249-260 (2004).

[2] R. Axelrod, W.D. Hamilton, The evolution of cooperation. *Science* **211**, 1390-1396 (1981).

[3] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[4] M. Broom, Evolutionary games with variable payoffs. *Comptes Rendus Biologies* **328**, 403-412 (2005).

[5] C. Cannings, The latent roots of certain Markov chains arising in genetics: A new approach, I. Haploid models. *Advances in Applied Probabilioty* **6**, 260-290 (1974).

[6] F.B. Christiansen, On conditions for evolutionary stability for a continuously varying character. *American Naturalist* **138**, 37-50 (1991).

[7] L.A. Dugatkin, *Cooperation Among Animals: An Evolutionary Perspective* (Oxford University Press, New York, 1997).

[8] I. Eshel and U. Motro, Kin selection and strong evolutionary stability of mutual help. *Theoretical Population Biology* **19**, 420-433 (1981).

[9] I. Eshel, L.L. Cavalli-Sforza, Assortment of encounters and evolution of cooperativeness. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1331-1335 (1982).

[10] I. Eshel, Evolutionary and continuous stability. *Journal of Theoretical Biology* **103**, 99-111 (1983).

[11] I. Eshel, On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution. *Journal of Mathematical Biology* **34**, 485-510 (1996).

[12] S.N. Ethier, T. Nagylaki. Diffusion approximations of Markov chains with two time scales and applications to population genetics. *Advances in Applied Probability* **12**, 14-49 (1980).

[13] W.J. Ewens, *Mathematical Population Genetics: I Theoretical Introduction* (Springer, New York, 2004).

[14] R.A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon, Oxford 1930).

[15] D. Fudenberg, D.G. Rand, A. Dreber, Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World. *American Economic Review* **102**(2), 720-749 (2012).

[16] T. Fujiwara-Greve, M. Okuno-Fujiwara, Voluntarily separable repeated prisoner's dilemma. *The Review of Economic Studies* **76**, 993-1021 (2009).

[17] W.D. Hamilton, The genetical evolution of social behaviour. *Journal of Theoretical Biology* **7**, 1-16 (1964).

[18] N. Hayashi, From TIT-for-TAT to OUT-for-TAT. *Sociol Theory Methods* **8**, 19-32 (1993).

[19] C. Hilbe, Š. Šimsa, K. Chatterjee, M.A. Nowak, Evolution of cooperation in stochastic games. *Nature* **559**, 246-249 (2018).

[20] J. Hofbauer, K. Sigmund, *The Theory of Evolution and Dynamical Systems.* (Cambridge University Press, Cambridge, UK, 1998).

[21] S.S. Izquierdo, L.R. Izquierdo, F. Vega-Redondo, The option to leave: Conditional dissociation in the evolution of cooperation. *Journal of Theoretical Biology* **267**, 76-84 (2010).

[22] S.S. Izquierdo, L.R. Izquierdo, F. Vega-Redondo, Leave and let leave: A sufficient condition to explain the evolutionary emergence of cooperation. *Journal of Economic Dynamics and Control* **46**, 91-113 (2014).

[23] H. Kaplan, K. Hill, A.M. Hurtado, "Risk, foraging and food sharing among the Ache" in *Risk and Uncertainty in Tribal and Peasant Economies*, pp. 107-144, E. Cashdan, Ed. (Westview Press, Boulder, CO, 1990).

[24] S. Karlin, U. Liberman, Random temporal variation in selection intensities: case of large population size. *Theoretical Population Biology* **6**, 355-382 (1974).

[25] S. Karlin, U. Liberman, Random temporal variation in selection intensities: One-locus two-allele model. *Journal of Mathematical Biology* **2**, 1-17 (1975).

[26] M. Kimura, Diffusion models in population genetics. *Journal of Applied Probability* **1**(2), 177-232 (1964).

[27] R. Lande, S. Engen, B.-E. Saether, *Stochastic Population Dynamics in Ecology and Conservation* (Oxford University Press, Oxford, 2003).

[28] S. Lessard, Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory. *Theoretical Population Biology* **68**, 19-27 (2005).

[29] S. Lessard, Effective Game Matrix and Inclusive Payoff in Group-Structured Populations. *Dynamic Games and Applications* **1**, 301-318 (2011).

[30] C. Li, B.-Y. Zhang, R. Cressman, Y. Tao, Evolution of cooperation in a heterogeneous graph: fixation probabilities under weak selection. *PLoS One* **8**, e66560 (2013).

[31] R.M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, 1973).

[32] R.M. May, Simple mathematical models with very complicated dynamics. *Nature* **261**, 459-467 (1976).

[33] J. Maynard Smith, G.R. Price, The logic of animal conflicts. *Nature* **246**, 15-18 (1973).

[34] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, UK, 1982).

[35] M.A. Nowak, R.M. May, Evolutionary games and spatial chaos. *Nature* **359**, 826-829 (1992).

[36] M.A. Nowak, K. Sigmund, A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* **364**, 56-58 (1993).

[37] M.A. Nowak, K. Sigmund, Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573-577 (1998a).

[38] M.A. Nowak, K. Sigmund, The Dynamics of Indirect Reciprocity. *Journal of Theoretical Biology* **194**, 561-574 (1998b).

[39] M.A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646-650 (2004).

[40] M.A. Nowak, K. Sigmund, Evolution of indirect reciprocity. *Nature* **437**, 1291-1298 (2005).

[41] M.A. Nowak, Five rules for the evolution of cooperation. *Science* **314**, 1560-1563 (2006a).

[42] M.A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006b).

[43] M.A. Nowak, K. Sigmund, "How populations cohere: Five rules for cooperation" in *Theoretical Ecology - Principles and Applications*, pp. 7-16, R.M. May, A. McLean, Eds. (Oxford University Press, Oxford, 2007).

[44] M.A. Nowak, R. Highfield, *Supercooperators: altruism, evolution, and why we need each other to succeed* (Free Press, New York, 2011).

[45] H. Ohtsuki, Y. Iwasa, How should we define goodness? Reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology* **231**, 107-120 (2004).

[46] H. Ohtsuki, C. Hauert, E. Lieberman, M.A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502-505 (2006).

[47] M. Perc, Coherence resonance in a spatial prisoner's dilemma game. *New Journal of Physics* **8**, 22-29 (2006).

[48] W. Poundstone, *Prisoner's Dilemma* (Doubleday, New York, 1992).

[49] R. Schuessler, Exit threats and cooperation under anonymity. *Journal of Conflict Resolution* **33**, 728-749 (1989).

[50] L.S. Shapley, Stochastic games. *Proc. Natl. Acad. Sci. U.S.A.* **39**(10), 1095-1100 (1953).

[51] K. Sigmund, The Calculus of Selfishness. (Princeton University Press, Princeton, 2010).

[52] E. Solan, N. Vieille, Stochastic games. *Proc. Natl. Acad. Sci. U.S.A.* **112**(45), 13743-13746 (2015).

[53] F. Stollmeier, J. Nagler, Unfair and anomalous evolutionary dynamics from fluctuating payoffs. *Physical Review Letters* **120**, 058101 (2018).

[54] A. Szolnoki, M. Perc, Seasonal payoff variations and the evolution of cooperation in social dilemmas. *Scientific Reports* **9**, 12575 (2019).

[55] Y. Tao, Q.-S. Yang, Z.-G. Jiang and Z.-W. Wang, Evolutionarily stable strategy, stable state, periodic cycle and chaos in a simple discrete time two-phenotype model. *Journal of Theoretical Biology* **188**, 21-27 (1997).

[56] P.D. Taylor, L. Jonker, Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* **40**, 145-156 (1978).

[57] A. Traulsen, M.A. Nowak, Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 10952-10955 (2006).

[58] R.L. Trivers, Evolution of reciprocal altruism. *Quarterly Review of Biology* **46**, 35-57 (1971).

[59] P. Uyttendaele, F. Thuijsman, P. Collins, R. Peeters, G. Schoenmakers, R. Westra, Evolutionary games and periodic fitness. *Dynamic Games and Applications* **2**, 335-345 (2012).

[60] S. Wright, Evolution in Mendelian populations. *Genetics* **16**, 97-159 (1931).

[61] E.C. Zeeman, in *Global Theory of Dynamical Systems* (Springer Verlag, Berlin, Heidelberg, 471, 1980).

[62] B.-Y. Zhang, C. Li, Y. Tao, Evolutionary Stability and the Evolution of Cooperation on Heterogeneous Graphs. *Dynamic Games and Applications* **6**, 567-579 (2016).

Part I

Part 1: Games in a stochastic environment

Chapter 1

Article 1

Evolutionary stability concepts in a stochastic environment

Xiu-Deng Zheng[1,*], Cong Li[2,*], Sabin Lessard[2] and Yi Tao[1,3]

[1]Key Laboratory of Animal Ecology and Conservation Biology, Centre for Computational and Evolutionary Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

[2]Department of Mathematics and Statistics, University of Montreal, Montreal QC H3C 3J7, Canada

[3]University of Chinese Academy of Sciences, Beijing 100101, China

[*]These authors contributed equally to this work.

Information

This paper appeared in *Physical Review E* **96**, 032414. It was received on 7 June 2017; published on 25 September 2017.

My contribution

I contributed to the early discussion for establishing the new concept of stochastic stability and the proofs of the theorems. Later I worked on the proofs of results 1.1 and 1.2, as well as all the numerical simulations and figures.

Abstract

Over the past thirty years, evolutionary game theory and the concept of an evolutionarily stable strategy have been not only extensively developed and successfully applied to explain the evolution of animal behaviors, but also widely used in economics and social sciences. Nonetheless, the stochastic dynamical properties of evolutionary games in randomly fluctuating environments are still unclear. In this study, we investigate conditions for stochastic local stability of fixation states and constant interior equilibria in a two-phenotype model with random payoffs following pairwise interactions. Based on this model, we develop the concepts of stochastic evolutionary stability (SES) and stochastic convergence stability (SCS). We show that the condition for a pure strategy to be SES and SCS is more stringent than in a constant environment, while the condition for a constant mixed strategy to be SES is less stringent than the condition to be SCS which is less stringent than the condition in a constant environment.

1.1 Introduction

Thirty four years ago, Maynard Smith's (1982) monograph "*Evolution and the Theory of Games*" was published. A new fundamental theoretical framework to understand the evolution of animal behavior had reached maturity and was finally made available to a large readership. Since then evolutionary game theory has been very popular not only in biology but also in economics and social sciences.

Evolutionary game theory started with the concept of *evolutionarily stable strategy* (ESS) introduced by Maynard Smith and Price (1973). Let us recall that an ESS is a strategy understood as a behavioural phenotype such that, if all the members of a population adopt it, then no mutant strategy could invade the population under the effect of natural selection (Maynard Smith, 1974, 1982). In the context of symmetric pairwise interactions occurring at random in an infinite population, a strategy $\mathbf{x}$ is an ESS if: (*i*) the payoff to $\mathbf{x}$ against itself is larger or equal to the payoff to any other strategy $\mathbf{y}$ against $\mathbf{x}$, and (*ii*) the payoff to $\mathbf{x}$ against $\mathbf{y}$ exceeds the payoff to $\mathbf{y}$ against itself in the case of an equality in (*i*). With $E(\mathbf{x},\mathbf{y})$ representing the payoff received by an individual using strategy $\mathbf{x}$ against an individual using strategy $\mathbf{y}$, this means that: (*i*) $E(\mathbf{x},\mathbf{x}) \geq E(\mathbf{y},\mathbf{x})$ for any strategy $\mathbf{y} \neq \mathbf{x}$, and (*ii*) $E(\mathbf{x},\mathbf{y}) > E(\mathbf{y},\mathbf{y})$ in the case of an equality in (*i*). These conditions are necessary and sufficient for the expected payoff to $\mathbf{x}$ to exceed the expected payoff to $\mathbf{y}$ in an infinite population of individuals using either $\mathbf{x}$ or $\mathbf{y}$ if the frequency of $\mathbf{y}$ is small enough.

If the relative growth rate of a strategy is given by its expected payoff, which defines its fitness, then the dynamics of the strategy frequencies is described by the *replicator equation* (Taylor and Jonker, 1978; Hofbauer and Sigmund, 1998). For n pure strategies, we have $\dot{x}_i = x_i\big((\mathbf{Ax})_i - \mathbf{xAx}\big)$, where $\mathbf{x} = (x_1, \ldots, x_n)$ is the strategy frequency vector, with x_i being the frequency of strategy i for $i = 1, \ldots, n$, and $\mathbf{A} = (a_{ij})$ is the payoff matrix, with a_{ij} being the payoff to strategy i against strategy j for $i, j = 1, \ldots, n$. Here, it is understood that $\mathbf{xAx} = \sum_{i=1}^{n} x_i(\mathbf{Ax})_i = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j a_{ij}$ with $(\mathbf{Ax})_i$ being the expected payoff to strategy i for $i = 1, \ldots, n$. Moreover, if $\mathbf{x}$ is an ESS with respect to the mixed strategies of the n pure strategies with the bilinear payoff function $E(\mathbf{x}, \mathbf{y}) = \mathbf{x}\mathbf{A}\mathbf{y}$, then it is an asymptotically stable rest point of the above replicator dynamics (Hofbauer and Sigmund, 1998). In the special case $n = 2$, the payoff matrix takes the form $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, and the replicator dynamics for the frequency of strategy 1, represented by x, reduces to $\dot{x} = x(1-x)\big((a_{11} - a_{12} - a_{21} + a_{22})x - (a_{22} - a_{21})\big)$. Moreover, strategy 1 is an ESS if $a_{11} > a_{21}$, or $a_{11} = a_{21}$ and $a_{12} > a_{22}$. These conditions are necessary and sufficient for the expected payoff to strategy 1 to exceed the expected payoff to strategy 2 in an infinite population when strategy 2 is rare enough.

In a population of fixed finite size N, any fixation state can be reached from any initial state by random drift. In this framework, Nowak *et al.* (2004) proposed to call strategy 1 an ESS$_N$ if two conditions hold when the initial frequency of strategy 2 is N^{-1}: (*i*) strategy 2 has a lower expected payoff than strategy 1 as in Schaffer (1988), in which case selection is said to oppose strategy 2 *invading* strategy 1; and (*ii*) the probability of ultimate fixation of strategy 2 is less than N^{-1}, in which case selection is said to oppose strategy 2 *replacing* strategy 1. In general, these conditions depend on the population size N and the reproduction scheme (Lessard and Ladret, 2007). Note that condition (*ii*) is neither sufficient nor necessary for the probability of ultimate fixation of a single strategy 1 to exceed the probability of ultimate fixation of a single strategy 2. This condition ensures, however, that strategy 1 is more *abundant* on average than strategy 2 in the presence of recurrent mutation occurring at weak enough rate (Antal *et al.*, 2009).

One key assumption in classical evolutionary game theory is that the payoff matrix is constant, and this supposes that the environmental conditions do not change over time. Previous work on stochastic evolutionary game theory in an infinite and classical population includes Foster and Young (1990) who considered small perturbations of the deterministic replicator dynamics that arise through mutations as well as ordinary chance events that affect the reproductive success of strategies. Then the strategy frequencies obey the stochastic differential equation $\dot{x}_i = x_i\left((\mathbf{A}\mathbf{x})_i - \mathbf{x}\mathbf{A}\mathbf{x}\right) + \sigma(\Gamma(\mathbf{x})\dot{\mathbf{W}})_i$. Here, $\dot{\mathbf{W}}$ is a formal time derivative of a standard n-dimensional Brownian motion $\mathbf{W}$, called a white noise, $\Gamma(\mathbf{x})$ is a variance-covariance matrix with all bounded entries and ones on the main diagonal such that $\mathbf{x}\Gamma(\mathbf{x}) = \mathbf{0}$, while $\sigma > 0$ is a parameter that represents the strength of the perturbation. In this stochastic dynamical system, a set of states S is called a *stochastically stable set* (SSS) if, in the long run, it is nearly certain that the system lies within every open set containing S as σ tends to zero. The stochastically stable set is always nonempty and minimizes a suitably defined potential function. However, it is by no means equivalent to the set of evolutionary stable strategies even when the latter exist. It contains often only a subset of the evolutionarily stable strategies, and sometimes even none. So, a natural and challenging question is what happens to evolutionary game concepts and dynamics under the effects of a stochastically varying environment.

As a matter of fact, environmental conditions in the real world are changing and uncertain. In turn, stochastic fluctuations in the surrounding environment of a population may cause changes in the occurrence of interactions between individuals and, more importantly, the payoffs received by the interacting individuals. Therefore, unless stochastic fluctuations are so small that their effects can be neglected, there is no a priori reason to assume that the payoff matrix of an evolutionary game is constant if the environment is actually stochastic. The role played by environmental fluctuations in the dynamics of biological and ecological systems has been investigated by a number of authors. These studies show the stochastic nature and the related noise induced effects in some population dynamics (Ciuchi *et al.*, 1993; Turchin *et al.*, 2000; Bjornstad and Grenfell, 2001; Ozbudak *et al.*, 2002; Elowitz *et al.*, 2002; Blake *et al.*, 2003; Paulsson, 2004; Spagnolo *et al.*, 2003, 2004; Chichigina *et al.*, 2005; Romanczuk *et al.*, 2009; La Cognata *et al.*, 2010; Chichigina *et al.*, 2011). However, the effects of environmental stochasticity on the evolutionary game dynamics and on some important concepts in evolutionary game theory (for example, the evolutionarily stable strategy) are still unclear.

Now assuming that the payoff matrix is random, two questions arise: First, how should we define the concept of *stochastic evolutionary stability* (SES) so that it would be a natural extension of the evolutionary stability concept in a stochastic environment in the sense that, once fixed, it would still probabilistically favored by selection? Second, what would be the exact evolutionary properties associated with a strategy that is SES? In particular, are there extra conditions that would make it *stochastically convergence stable* (SCS) in such a way that evolution toward it from other fixation states would be probabilistically favored by selection under random perturbations in an analogous way as a convergence stable strategy is favored in a deterministic environment (Eshel and Motro, 1981; Eshel, 1983, 1996; Christiansen, 1991). Answers to these questions are important in order to understand and predict the evolution of animal behaviors in a randomly fluctuating environment.

In this study, we focus attention on the effect of a stochastic environment on a 2×2 matrix game in an infinite population. Generations are discrete, nonoverlapping, and the payoff matrices over successive generations are independent identically distributed random matrices. The main mathematical tool in this study is the concept of *stochastic local stability*, which was developed in population genetics by Karlin and Liberman (1974, 1975) (see also Ewens 2004).

1.2 A Two-Phenotype Model

We consider an evolutionary game in an infinite population with discrete, nonoverlapping, generations. There are two phenotypes or pure strategies, 1 and 2, and the payoffs in pairwise interactions at time step $t \geq 0$ are given by the game matrix

$$\mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix} = \begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix}, \tag{1.2.1}$$

where $a_{ij}(t)$ is the payoff to strategy i against strategy j for $i, j = 1, 2$. These payoffs are assumed to be positive random variables that are uniformly bounded below and above by some positive constants. Therefore, there exist real numbers $A, B > 0$ such that $A \leq a_{ij}(t) \leq B$ for $i, j = 1, 2$ and all $t \geq 0$. Moreover, the probability distributions of $a_{ij}(t)$ for $i, j = 1, 2$ do not depend on $t \geq 0$. They have means, variances and covariances given by

$$\begin{aligned} \mathbb{E}\left(a_{ij}(t)\right) &= \bar{a}_{ij} , \\ \mathbb{E}\left((a_{ij}(t) - \bar{a}_{ij})^2\right) &= \sigma_{ij}^2 , \\ \mathbb{E}\big((a_{ij}(t) - \bar{a}_{ij})(a_{kl}(t) - \bar{a}_{kl})\big) &= \sigma_{ij,kl} , \end{aligned} \tag{1.2.2}$$

respectively, for $i, j, k, l = 1, 2$ with $(i, j) \neq (k, l)$, where $\mathbb{E}$ denotes mathematical expectation. As for $s \neq t$, the payoffs $a_{ij}(s)$ and $a_{kl}(t)$ are assumed to be independent so that $\mathbb{E}\big((a_{ij}(s) - \bar{a}_{ij})(a_{kl}(t) - \bar{a}_{kl})\big) = 0$ for $i, j, k, l = 1, 2$.

Let x_t be the frequency of strategy 1 at time step $t \geq 0$ and, similarly, $1 - x_t$ the frequency of strategy 2. Then the mean payoffs to strategies 1 and 2 are given by

$$\begin{aligned} \pi_{1,t} &= x_t a_t + (1 - x_t) b_t , \\ \pi_{2,t} &= x_t c_t + (1 - x_t) d_t , \end{aligned} \tag{1.2.3}$$

respectively, and the mean payoff in the whole population by

$$\bar{\pi}_t = x_t \pi_{1,t} + (1 - x_t) \pi_{2,t} . \tag{1.2.4}$$

Assuming that payoff translates into reproductive success, referred to as fitness, so that the number of replicas of a strategy from one step to the next is proportional to its mean payoff, the frequency of strategy 1 at time step $t + 1$ is given by the recurrence equation

$$x_{t+1} = \frac{x_t \pi_{1,t}}{\bar{\pi}_t} = \frac{x_t^2 a_t + x_t(1 - x_t) b_t}{x_t^2 a_t + x_t(1 - x_t)(b_t + c_t) + (1 - x_t)^2 d_t} \tag{1.2.5}$$

for $t \geq 0$. Defining

$$u_t = \frac{x_t}{1 - x_t} , \tag{1.2.6}$$

the recurrence equation takes the simple form

$$u_{t+1} \quad = \quad u_t \left[\frac{u_t a_t + b_t}{u_t c_t + d_t} \right] \tag{1.2.7}$$

for $t \geq 0$.

1.3 Stochastic Local Stability

We are interested in the asymptotic (or long run) behavior of the process $\{x_t\}$ for $t \geq 0$. Let $\hat{x}$ represent a constant (non-random) equilibrium of this process, that is, an equilibrium of Eq. (1.2.5) that does not depend on the randomness of the payoff matrix. This is clearly the case for both $\hat{x} = 0$ and $\hat{x} = 1$, called the *fixation states* or the *boundary equilibria*. This may also be the case for a constant equilibrium $\hat{x}$ with $0 < \hat{x} < 1$, called a *constant interior equilibrium*.

Following Karlin and Liberman (1974, 1975; Ewens, 2004), a constant equilibrium $\hat{x}$ is said to be *stochastically locally stable* (SLS) if for any $\epsilon > 0$ there exists $\delta_0 > 0$ such that

$$\mathbb{P}\left(x_t \rightarrow \hat{x}\right) \geq 1 - \epsilon \quad \text{as soon as} \quad |x_0 - \hat{x}| < \delta_0 \ . \tag{1.3.1}$$

This means that x_t tends to $\hat{x}$ as $t \rightarrow \infty$ with probability arbitrarily close to 1 (but different from 1) if the initial state x_0 is sufficiently near $\hat{x}$. Notice, however, no matter how close x_0 is to $\hat{x}$ (but different from $\hat{x}$), it is not ascertained that x_t will converge to $\hat{x}$. Statistical fluctuations could cause x_t to depart sharply from $\hat{x}$, but this will occur with small probability if x_0 is close to $\hat{x}$ and $\hat{x}$ is stochastically locally stable.

On the other hand, a constant equilibrium $\hat{x}$ can be said to be *stochastically locally unstable* (SLU) if

$$\mathbb{P}\left(x_t \rightarrow \hat{x}\right) = 0 \quad \text{as soon as} \quad |x_0 - \hat{x}| > 0 \ . \tag{1.3.2}$$

If this is the case, then $\hat{x}$ cannot be reached with probability 1 from any initial state different from $\hat{x}$.

1.3.1 Stochastic local stability of fixation states

Consider first the fixation state $\hat{x} = 0$ in Eq. (1.2.5), which corresponds to the equilibrium $\hat{u} = \hat{x}/(1 - \hat{x}) = 0$ in Eq. (1.2.7). The sufficient condition for this equilibrium to be SLS is

$$\mathbb{E}\left(\log\left(\frac{d_t}{b_t}\right)\right) \;=\; \mathbb{E}\left(\log d_t\right) - \mathbb{E}\left(\log b_t\right) \;>\; 0 \tag{1.3.3}$$

(see details in **Appendix A. Theorem 1**). Therefore, under generic conditions, the inequality Eq. (1.3.3) is necessary and sufficient for stochastic local stability of the fixation state $\hat{x} = 0$. This result in a population genetics framework, which corresponds to a symmetric game matrix ($a_{12}(t) = a_{21}(t)$ which can be assumed equal to one without loss of generality) was stated in Karlin and Liberman (1974). A proof in this framework which only slightly differs from the present more general game-theoretic framework is given in Karlin and Liberman (1975). It is based on the strong law of large numbers and Egorov's theorem.

Suppose random payoffs close enough to their means so that

$$\begin{aligned}\mathbb{E}\left(\log b_t\right) &= \log \bar{b} - \frac{\sigma_b^2}{2\bar{b}^2} + o\left(\sigma^2\right) \ , \\ \mathbb{E}\left(\log d_t\right) &= \log \bar{d} - \frac{\sigma_d^2}{2\bar{d}^2} + o\left(\sigma^2\right) \ , \end{aligned} \tag{1.3.4}$$

where $\bar{b}$ and $\bar{d}$ are the means, and σ_b^2 and σ_d^2 the variances, of the random variables b_t and d_t, respectively. Here, $\sigma^2 = \max\{\sigma_a^2, \sigma_b^2, \sigma_c^2, \sigma_d^2\}$. Thus, if σ^2 is small enough, the condition in Theorem 1 for $\hat{x} = 0$ to be SLS reduces to

$$\log\left(\frac{\bar{d}}{\bar{b}}\right) \;>\; \frac{1}{2}\left(\frac{\sigma_d^2}{\bar{d}^2} - \frac{\sigma_b^2}{\bar{b}^2}\right) \ . \tag{1.3.5}$$

If the inequality is reversed, then $\hat{x} = 0$ is SLU. Therefore, the condition for $\hat{x} = 0$ to be SLS becomes less stringent as σ_b^2 increases and more stringent as σ_d^2 decreases. In the case where $\sigma_b^2 = \sigma_d^2 = \sigma^2$, the fixation state $\hat{x} = 0$ is SLS if $\bar{d} > \bar{b}$ and SLU if $\bar{d} < \bar{b}$. These are the conditions for $\hat{x} = 0$ to be locally stable and locally unstable, respectively, with a constant payoff matrix which corresponds to the case $\sigma^2 = 0$.

By symmetry, Theorem 1.1 implies that the fixation state $\hat{x} = 1$ in the recurrence equation (Eq. (1.2.5)) is stochastically locally stable if

$$\mathbb{E}\left(\log\left(\frac{a_t}{c_t}\right)\right) > 0 \ , \tag{1.3.6}$$

and stochastically locally unstable if the inequality is reversed. The above condition reduces to

$$\log\left(\frac{\bar{a}}{\bar{c}}\right) > \frac{1}{2}\left(\frac{\sigma_a^2}{\bar{a}^2} - \frac{\sigma_c^2}{\bar{c}^2}\right) \tag{1.3.7}$$

if a_t and c_t have means $\bar{a}$ and $\bar{c}$, and variances σ_a^2 and σ_c^2 of order σ^2 small enough.

As an example, consider successive rounds of the Prisoner's Dilemma, known as the iterated Prisoner's Dilemma (IPD), with two possible strategies in use, TFT for *tit-for-tat* starting with cooperation in the first round and the previous strategy of the opponent in the next rounds and AllD for *always-defect* with defection in all rounds as strategies 1 and 2, respectively (Axelrod, 1984; Nowak, 2006). The payoff matrix at time step $t \geq 0$ is given by

$$\begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} = \begin{pmatrix} m_t R & S + (m_t - 1)P \\ T + (m_t - 1)P & m_t P \end{pmatrix} , \tag{1.3.8}$$

where $T > R > P > S$ are the payoffs in one round of the game with cooperation and defection as strategies, and m_t represents the number of rounds at time step $t \geq 0$. This number is assumed to be a random variable of mean $\bar{m}$ and small variance σ_m^2 that is independent of m_s for all $s \neq t$. We are in the above context with

$$\bar{a} = \bar{m}R, \quad \bar{b} = S + (\bar{m} - 1)P, \quad \bar{c} = T + (\bar{m} - 1)P, \quad \bar{d} = \bar{m}P, \tag{1.3.9}$$

while

$$\sigma_a^2 = R^2\sigma_m^2, \quad \sigma_b^2 = \sigma_c^2 = \sigma_d^2 = P^2\sigma_m^2. \tag{1.3.10}$$

Since $\sigma_b^2 = \sigma_d^2$ and $\bar{d} > \bar{b}$, AllD-fixation is SLS. On the other hand, TFT-fixation is SLS if

$$\log\left(\frac{\bar{a}}{\bar{c}}\right) \quad > \quad \frac{\sigma_m^2}{2\bar{c}^2\bar{a}^2}\left(\bar{c}^2R^2 - \bar{a}^2P^2\right) \tag{1.3.11}$$

with

$$-\bar{c}^2R^2 - \bar{a}^2P^2 = \big(\bar{m}RP + (T-R)R\big)^2 - (\bar{m}RP)^2 > 0 \ . \tag{1.3.12}$$

This means a more stringent condition as σ_m^2 increases compared to the condition $\bar{a} > \bar{c}$ when $\sigma_m^2 = 0$, which is equivalent to $\bar{m} > (T-P)/(R-P)$. In a stochastic environment, the mean number of rounds must exceed a higher threshold value for TFT-fixation to be SLS. For instance, let $T = 4$, $R = 3$, $P = 2$ and $S = 1$. When $\sigma_m^2 = 0$, TFT-fixation is SLS if $\bar{m} > 2$. When $\sigma_m^2 > 0$, however, this occurs if

$$\log\left(\frac{3\bar{m}}{2(\bar{m}+1)}\right) > \frac{\sigma_m^2}{2\bar{m}^2(\bar{m}+1)^2}(2\bar{m}+1) \ . \tag{1.3.13}$$

This inequality is satisfied if and only if $\bar{m} > m(\sigma_m^2)$, where the threshold value $m(\sigma_m^2)$ is an increasing function of σ_m^2 (see Fig. 1.1). This clearly illustrates the fact that stochastic local stability of TFT-fixation depends not only on the mean of the number of rounds $\bar{m}$ but also on its variance σ_m^2, and that higher is the variance, higher must be the mean for TFT-fixation to be SLS.

Returning to Theorem 1.1, the stochastic local stability or instability of the fixation state $\hat{x} = 0$ in the degenerate case $b_t = d_t$ for all $t \geq 0$ requires further analysis as presented in **Appendix B. Theorem 2** .

Developing the random variables around their means and using the approximations

$$\begin{aligned} \mathbb{E}\left(\frac{a_t}{d_t}\right) &= \frac{\bar{a}}{\bar{d}} + \frac{\bar{a}\sigma_d^2}{\bar{d}^3} - \frac{\sigma_{a,d}}{\bar{d}^2} + o\left(\sigma^2\right) \ , \\ \mathbb{E}\left(\frac{c_t}{d_t}\right) &= \frac{\bar{c}}{\bar{d}} + \frac{\bar{c}\sigma_d^2}{\bar{d}^3} - \frac{\sigma_{c,d}}{\bar{d}^2} + o\left(\sigma^2\right) \ , \end{aligned} \tag{1.3.14}$$

the condition in Theorem 2 for $\hat{x} = 0$ to be SLS reduces to

$$\frac{\bar{c} - \bar{a}}{\bar{d}} > \frac{\sigma_{c,d} - \sigma_{a,d}}{\bar{d}^2 + \sigma_d^2} \tag{1.3.15}$$

if σ^2 is small enough. If the inequality is reversed, then $\hat{x} = 0$ is SLU. Therefore, the condition for $\hat{x} = 0$ to be SLS becomes less stringent as $\sigma_{a,d}$ increases and more stringent as $\sigma_{c,d}$ decreases. In the case where $\sigma_{a,d} = \sigma_{c,d}$, the fixation state $\hat{x} = 0$ is SLS if $\bar{c} > \bar{a}$ and

SLU if $\bar{c} < \bar{a}$. These are the conditions for $\hat{x} = 0$ to be locally stable and locally unstable, respectively, with a constant payoff matrix.

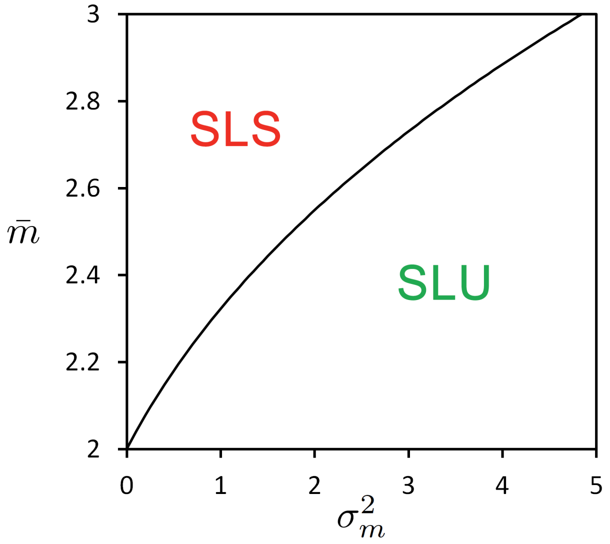


Figure 1.1: *Stochastic local stability in IPD game.* Stochastic local stability of TFT-fixation against AllD with the random payoff matrix $\begin{pmatrix} 3m_t & 2m_t - 1 \\ 2(m_t+1) & 2m_t \end{pmatrix}$, where $T = 4$, $R = 3$, $P = 2$ and $S = 1$, while m_t is a random variable with mean $\bar{m}$ and variance σ_m^2 at every time step $t \geq 0$. The curve separates the regions for stochastic local stability and stochastic local instability: TFT-fixation is SLS (or SLU) if the point $\left(\sigma_m^2, \bar{m}\right)$ is above (or below) the curve.

1.3.2 Stochastic local stability of a constant interior equilibrium

Now consider a constant interior equilibrium $\hat{x}$ of Eq. (1.2.5) with $0 < \hat{x} < 1$. This corresponds to a constant equilibrium $\hat{u} = \hat{x}/(1-\hat{x}) > 0$ in Eq. (1.2.7). This is possible only if

$$\hat{u}(a_t - c_t) \quad = \quad d_t - b_t \;, \tag{1.3.16}$$

which implies a payoff matrix in the form

$$\begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} = \begin{pmatrix} c_t + z_t & b_t \\ c_t & b_t + \hat{u} z_t \end{pmatrix} = \begin{pmatrix} a_t & d_t - \hat{u} z_t \\ a_t - z_t & d_t \end{pmatrix} , \qquad (1.3.17)$$

where $z_t = a_t - c_t$. Moreover, taking expectation on both sides of Eq. (1.3.16) yields $\hat{u} = \left(\bar{d} - \bar{b}\right) / \left(\bar{a} - \bar{c}\right)$ from which $\hat{x} = \left(\bar{d} - \bar{b}\right) / \left(\bar{a} - \bar{b} - \bar{c} + \bar{d}\right)$, where $\bar{a}$, $\bar{b}$, $\bar{c}$ and $\bar{d}$ denote the expected values of a_t, b_t, c_t and d_t, respectively.

The sufficient condition for this constant equilibrium $\hat{x}$ with $\hat{u} > 0$ to be SLS is

$$\mathbb{E}\left(\log\left(\frac{\hat{u} c_t + d_t}{\hat{u} a_t + d_t}\right)\right) = \mathbb{E}\left(\log\left(1 - \frac{\hat{x} z_t}{\hat{x} a_t + (1 - \hat{x}) d_t}\right)\right) > 0 , \qquad (1.3.18)$$

(see details in **Appendix C. Theorem 3**).

Assuming the approximation

$$\mathbb{E}\left(\log\left(\hat{u} a_t + d_t\right)\right) = \log\left(\hat{u}\bar{a} + \bar{d}\right) - \frac{\hat{u}^2 \sigma_a^2 + \sigma_d^2 + 2\hat{u}\sigma_{a,d}}{2\left(\hat{u}\bar{a} + \bar{d}\right)^2} + o\left(\sigma^2\right) \qquad (1.3.19)$$

and the corresponding approximation for $\mathbb{E}\left(\log\left(\hat{u} c_t + d_t\right)\right)$ lead to the condition

$$\log\left(\frac{\hat{u}\bar{c} + \bar{d}}{\hat{u}\bar{a} + \bar{d}}\right) > \frac{1}{2}\left(\frac{\hat{u}^2 \sigma_c^2 + 2\hat{u}\sigma_{c,d} + \sigma_d^2}{\left(\hat{u}\bar{c} + \bar{d}\right)^2} - \frac{\hat{u}^2 \sigma_a^2 + 2\hat{u}\sigma_{a,d} + \sigma_d^2}{\left(\hat{u}\bar{a} + \bar{d}\right)^2}\right) \qquad (1.3.20)$$

for the equilibrium $\hat{x} = \hat{u}/(1 + \hat{u})$ to be SLS if σ^2 is small enough. The reversed inequality guarantees that $\hat{x}$ is SLU. If a_t, d_t and z_t are independent random variables, then the above condition takes the form

$$\log\left(1 - \frac{\hat{x}\bar{z}}{\hat{x}\bar{a} + (1 - \hat{x})\bar{d}}\right) > \frac{1}{2}\left(\frac{\hat{u}^2 \sigma_a^2 + \hat{u}^2 \sigma_z^2 + \sigma_d^2}{\left(\hat{u}\bar{a} - \hat{u}\bar{z} + \bar{d}\right)^2} - \frac{\hat{u}^2 \sigma_a^2 + \sigma_d^2}{\left(\hat{u}\bar{a} + \bar{d}\right)^2}\right), \qquad (1.3.21)$$

where $\bar{z}$ and σ_z^2 designate the mean and variance of z_t, respectively. When all variances vanish, the condition reduces to $\bar{z} < 0$, which means that $\bar{a} < \bar{c}$ and $\bar{d} < \bar{b}$. Notice that this condition becomes more stringent as σ_z^2 increases.

It is worth emphasizing that it is possible for a constant interior equilibrium and both fixation states to be simultaneously SLS. For instance, consider a payoff matrix in the form

$$\begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} = \begin{pmatrix} 1 & 1 + \hat{u}\eta_t \\ 1 + \eta_t & 1 \end{pmatrix} , \qquad (1.3.22)$$

where $\hat{u} > 0$ and $\eta_t = -z_t$ is a random variable with mean $\bar{\eta} = -\bar{z} > 0$ and variance $\sigma_\eta^2 = \sigma_z^2$. Then, it is easy to check that: (i) the fixation state $\hat{x} = 0$ is SLS if

$$\sigma_\eta^2 > 2\left(\frac{1+\hat{u}\bar{\eta}}{\hat{u}}\right)^2 \log(1+\hat{u}\bar{\eta}) \ ; \tag{1.3.23}$$

(ii) the fixation state $\hat{x} = 1$ is SLS if

$$\sigma_\eta^2 > 2(1+\bar{\eta})^2 \log(1+\bar{\eta}) \ ; \tag{1.3.24}$$

and (iii) the constant interior equilibrium $\hat{x} = \hat{u}/(1+\hat{u})$ is SLS if

$$\sigma_\eta^2 \ < \ 2\left(\frac{1+\hat{x}\bar{\eta}}{\hat{x}}\right)^2 \log\left(1+\hat{x}\bar{\eta}\right) \ . \tag{1.3.25}$$

On the other hand, the three equilibria are SLU if all the inequalities are reversed. In the special case where $\hat{u} = 1$ (i.e., $\hat{x} = 1/2$), for instance, it can be shown that there exists a threshold value $\eta_0 > 0$ such that $(1+\bar{\eta})^2 \log(1+\bar{\eta}) < (2+\bar{\eta})^2 \log\big(1+\bar{\eta}/2\big)$ if and only if $\bar{\eta} < \eta_0$. Therefore, the constant interior equilibrium $\hat{x} = 1/2$ and both fixation states, $\hat{x} = 0$ and $\hat{x} = 1$, are simultaneously SLS when $0 < \bar{\eta} < \eta_0$ and

$$2(1+\bar{\eta})^2 \log(1+\bar{\eta}) < \sigma_\eta^2 < 2(2+\bar{\eta})^2 \log\big(1+\bar{\eta}/2\big) \ . \tag{1.3.26}$$

On the other hand, the three equilibria are simultaneously SLU when all the inequalities are reversed. The corresponding regions for the different regimes are illustrated in Fig. 1.2a and simulation results with a constant interior equilibrium and both fixation states simultaneously SLS are presented in Fig. 1.2b.

1.4 Stochastic Evolutionary Stability and Stochastic Convergence Stability

Extending the standard definition of an evolutionarily stable strategy (ESS) in a constant environment (Maynard Smith and Price, 1973) to a variable environment, a *stochastically*

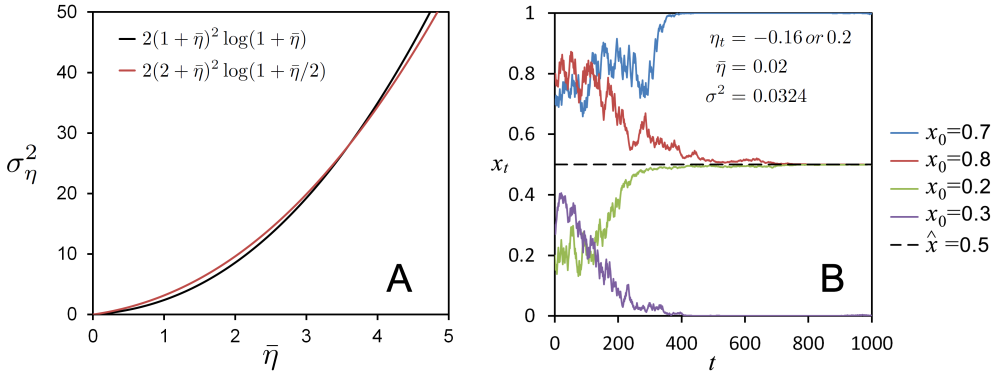


Figure 1.2: *Stochastic local stability of a constant interior equilibrium and of both fixation states.* Stochastic local stability or instability of a constant interior equilibrium and of both fixation states with the random payoff matrix $\begin{pmatrix} 1 & 1+\eta_t \\ 1+\eta_t & 1 \end{pmatrix}$. (**A**) The black curve represents the function $\sigma_\eta^2 = 2(1+\bar{\eta})^2\log(1+\bar{\eta})$ and the red curve the function $\sigma_\eta^2 = 2(2+\bar{\eta})^2\log(1+\bar{\eta}/2)$. There is a critical value of $\bar{\eta}$, denoted by η_0, that corresponds to the intersection of the black and red curves. For $\bar{\eta} < \eta_0$, all of $\hat{x} = 0$, $\hat{x} = 1$ and $\hat{x} = 1/2$ are SLS if the the point $\left(\sigma_\eta^2, \bar{\eta}\right)$ is in the range between the black and red curves; and, conversely, for $\bar{\eta} > \eta_0$, all of $\hat{x} = 0$, $\hat{x} = 1$ and $\hat{x} = 1/2$ are SLU if the point $\left(\sigma_\eta^2, \bar{\eta}\right)$ is in the range between the black and red curves. (**B**) The simulation results, where $\hat{u} = 1$ and $\eta_t = -0.16$ and 0.2 with same probability 0.5 so that $\bar{\eta} = 0.02$ and $\sigma_\eta^2 = 0.0324$. Four trajectories of x_t, the frequency of strategy 1, are illustrated starting with $x_0 = 0.2, 0.3, 0.7, 0.8$: two converge to $\hat{x} = 1/2$, one to $\hat{x} = 0$, and one to $\hat{x} = 1$. Here, three equilibrium states $\hat{x} = 0$, $\hat{x} = 1$ and $\hat{x} = 1/2$ are SLS.

evolutionarily stable (SES) strategy can be defined as a strategy such that, if all the members of the population adopt it, then the probability for at least any slightly perturbed strategy to invade the population under the influence of natural selection is arbitrarily low. More specifically, a strategy represented by a frequency vector $\hat{\mathbf{x}}$ is SES if $\hat{\mathbf{x}}$-fixation is SLS against any other strategy $\mathbf{x} \neq \hat{\mathbf{x}}$ at least near enough.

Similarly, the notion of a continuous stable strategy (CSS) introduced in Eshel and Motro (1981) (see also Eshel, 1983) and renamed later on convergence stable strategy (Chrstiansen, 1991; Eshel, 1996) can be extended to a context of a variable environment. So, a strategy represented by a two-dimensional frequency vector $\hat{\mathbf{x}} = (\hat{x}, 1-\hat{x})$ can be said to be *stochastically convergence stable* (SCS) if the fixation state of any nearby strategy $\tilde{\mathbf{x}} = (\tilde{x}, 1-\tilde{x})$ is SLU against a strategy $\mathbf{x} = (x, 1-x)$ if and only if $\mathbf{x}$ is in the direction of $\hat{\mathbf{x}}$ with respect to $\tilde{\mathbf{x}}$. This means that $\text{sgn}(x-\tilde{x}) = \text{sgn}(\hat{x}-\tilde{x})$.

In this section, assume a positive stochastic game matrix at each time step $t \geq 0$ in the form

$$\begin{aligned}
\mathbf{A}(t) &= \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix} \\
&= \begin{pmatrix} \bar{a}_{11} & \bar{a}_{12} \\ \bar{a}_{21} & \bar{a}_{22} \end{pmatrix} + \begin{pmatrix} b_{11}(t) & b_{12}(t) \\ b_{21}(t) & b_{22}(t) \end{pmatrix} \\
&= \bar{\mathbf{A}} + \mathbf{B}(t) \ . \qquad (1.4.1)
\end{aligned}$$

Here, $\bar{\mathbf{A}}$ is a constant matrix with all positive entries, while the entries of $\mathbf{B}(t)$ are stochastic with mean equal to 0. The payoff matrix at time step $t \geq 0$ for two mixed strategies, $\mathbf{x} = (x, 1-x)$ and $\hat{\mathbf{x}} = (\hat{x}, 1-\hat{x})$ in this order, is then given by

$$\begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} = \begin{pmatrix} \mathbf{x}\mathbf{A}(t)\mathbf{x} & \mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}} \\ \hat{\mathbf{x}}\mathbf{A}(t)\mathbf{x} & \hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}} \end{pmatrix} \ . \qquad (1.4.2)$$

We are now ready to state our next two results which are proved in **Appendix D and E**, respectively.

Result 1.1: Consider the positive stochastic game matrix Eq. (1.4.1) where $b_{ij}(t)$ for $i, j = 1, 2$ are independent random variables of variances σ_{ij}^2 for $i, j = 1, 2$, respectively, while all higher-order centered moments are functions $o(\sigma^2)$ where $\sigma^2 = \max\{\sigma_{ij}^2 : i, j = 1, 2\}$. With the assumption that σ^2 is small enough and under generic conditions, the pure strategy $\hat{\mathbf{x}} = (0, 1)$ is stochastically evolutionarily stable against any nearby mixed strategy $\mathbf{x} = (x, 1-x)$ if and only if

$$\sigma_{22}^2 \quad < \quad \left(\frac{\bar{a}_{22}}{\bar{a}_{12}} - 1\right)\bar{a}_{22}^2 \ . \tag{1.4.3}$$

This is also the necessary and sufficient condition for $\hat{\mathbf{x}} = (0, 1)$ to be stochastically convergence stable. By symmetry, the pure strategy $\hat{\mathbf{x}} = (1, 0)$ is stochastically evolutionarily stable against any nearby mixed strategy $\mathbf{x} = (x, 1-x)$ and stochastically convergence stable in generic cases under the condition that σ^2 is small enough if and only if

$$\sigma_{11}^2 \quad < \quad \left(\frac{\bar{a}_{11}}{\bar{a}_{21}} - 1\right)\bar{a}_{11}^2 \ . \tag{1.4.4}$$

Result 1.2: Consider the positive stochastic game matrix Eq. (1.4.1) where $a_{11}(t) = a_{21}(t) + z(t)$ and $a_{22}(t) = a_{12}(t) + \hat{u}z(t)$ for some constant $\hat{u} = \hat{x}/(1-\hat{x}) > 0$, with $a_{21}(t)$, $a_{12}(t)$ and $z(t)$ being independent random variables of means and variances given by $\bar{a}_{21}$, $\bar{a}_{12}$, $\bar{z}$ and σ_{21}^2, σ_{12}^2, σ_z^2, respectively, while all higher-order centered moments are functions $o(\sigma^2)$, where $\sigma^2 = \max\{\sigma_{21}^2, \sigma_{12}^2, \sigma_z^2\}$. With the assumption that σ^2 is small enough and under generic conditions, the constant mixed strategy $\hat{\mathbf{x}} = (\hat{x}, 1-\hat{x})$ is stochastically evolutionarily stable against any mixed strategy $\mathbf{x} = (x, 1-x)$ if and only if

$$\bar{z}\bar{d}^2 + \bar{z}\sigma_{21}^2\hat{x}^2 + \bar{z}\sigma_{12}^2(1-\hat{x})^2 - \bar{d}\sigma_z^2\hat{x} < 0, \tag{1.4.5}$$

where

$$\bar{d} = \bar{z}\hat{x} + \bar{a}_{21}\hat{x} + \bar{a}_{12}(1-\hat{x}) > 0. \tag{1.4.6}$$

On the other hand, it is stochastically convergence stable under the same assumption and conditions if and only if

$$\bar{z}\bar{d}^2 + \bar{z}\sigma_{21}^2\hat{x}^2 + \bar{z}\sigma_{12}^2(1-\hat{x})^2 + \bar{z}\hat{x}^2\sigma_z^2 - \bar{d}\sigma_z^2\hat{x} < 0 \ . \qquad (1.4.7)$$

Consider, for instance, a positive stochastic game matrix

$$\mathbf{A}(t) \quad = \quad \begin{pmatrix} 1+z(t) & 1 \\ 1 & 1+\hat{u}z(t) \end{pmatrix} \ , \qquad (1.4.8)$$

where $\hat{u} = \hat{x}/(1-\hat{x}) > 0$ and $z(t)$ is a random variable of mean $\bar{z}$ and variance σ_z^2 such that $\bar{d} = \bar{z}\hat{x} + 1 > 0$ and σ_z^2 is small enough. Owing to **Result 1.1**, conditions Eqs. (1.4.3-1.4.4) for the pure strategies $(0,1)$ and $(1,0)$ to be SES and SCS are

$$\hat{u}\sigma_z^2 < \bar{z}\left(1+\hat{u}\bar{z}\right)^2 \qquad (1.4.9)$$

and

$$\sigma_z^2 < \bar{z}\left(1+\bar{z}\right)^2 \ , \qquad (1.4.10)$$

respectively. On the other hand, conditions Eqs. (1.4.5, 1.4.7) for the constant mixed strategy $\hat{\mathbf{x}} = (\hat{x}, 1-\hat{x})$ where $\hat{x} = \hat{u}/(1+\hat{u})$ to be SES and SCS reduce to

$$\hat{x}\sigma_z^2 > \bar{z}\left(1+\hat{x}\bar{z}\right) \qquad (1.4.11)$$

and

$$\hat{x}\sigma_z^2 > \bar{z}\left(1+\hat{x}\bar{z}\right)^2 \ , \qquad (1.4.12)$$

respectively. Figure 1.3 illustrates the corresponding regions in the case $\hat{u} = 1$ which corresponds to $\hat{x} = 1/2$. In the limit of a deterministic game matrix, that is, $\sigma_z^2 = 0$, both pure strategies are SES and SCS when $\bar{z} > 0$, while the mixed strategy is SES and SCS when $\bar{z} < 0$. In the presence of stochastic perturbations on the game matrix, that is, $\sigma_z^2 = \sigma^2 > 0$ (where σ^2 is small), there exist three threshold values $z_1 > z_2 > z_3 > 0$ such that both pure strategies are SES and SCS when $\bar{z} > z_1$ (where z_1 is a positive solution of $\sigma_z^2 = \bar{z}(1+\bar{z})^2$), while the mixed strategy is SES when $\bar{z} < z_2$ (where z_2 is a positive solution of $\sigma_z^2 = 2\bar{z}(1+\bar{z}/2)$) and SCS when $\bar{z} < z_3$ (where z_3 is a positive solution of $\sigma_z^2 = 2\bar{z}(1+\bar{z}/2)^2$).

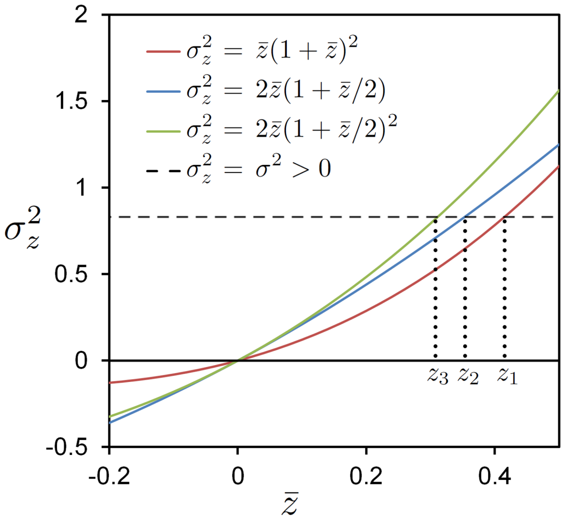


Figure 1.3: *Stochastic evolutionary stability and stochastic convergence stability.* For the positive stochastic payoff matrix $\begin{pmatrix} 1+z(t) & 1 \\ 1 & 1+z(t) \end{pmatrix}$ with a constant interior equilibrium $\hat{x} = 1/2$, if σ_z^2 is small, then both pure strategies $(0,1)$ and $(1,0)$ are SES and SCS when $\bar{z} > z_1$, and the constant mixed strategy $(1/2, 1/2)$ is SES when $\bar{z} < z_2$ and SCS when $\bar{z} < z_3$.

1.5 Conclusion

Evolutionary concepts such as that of an evolutionarily stable strategy (ESS) (Maynard Smith and Price, 1973) and that of a convergence stable strategy (CSS) (Eshel and Motro, 1981; Christiansen, 1991) were originally introduced for infinite populations in a deterministic environment. Therefore, they were initially stated in terms of conditions that ensure local (actually, asymptotic) stability of a resident strategy against any mutant strategy, or local instability (actually, initial invasion) of any resident strategy close enough to a given population strategy following the introduction of any mutant that brings the population strategy even closer.

In a stochastic environment, convergence to a constant equilibrium from any given initial state occurs with some probability. When this probability tends to 1 as the initial state tends to the equilibrium, then the equilibrium is said to be stochastically locally stable (SLS). On the other hand, when this probability is always 0 for any initial state different from the equilibrium, then the equilibrium is said to be stochastically locally unstable (SLU). These conditions were studied by Karlin and Liberman (1974, 1975) in the framework of a one-locus two-allele viability model for a random mating diploid population undergoing discrete, nonoverlapping generations. This framework corresponds to a linear game model in discrete time with a symmetric payoff matrix for two pure strategies used in random pairwise interactions. We have extended the analysis to a general payoff matrix.

In the absence of stochastic perturbations on the payoffs, it is well known that a fixed resident strategy is locally stable against a mutant strategy introduced in small frequency if the payoff of the resident strategy against itself exceeds the payoff of the mutant strategy against the resident strategy, or in case of equality, if the payoff of the resident strategy against the mutant strategy exceeds the payoff of the mutant strategy against itself. In the presence of stochastic perturbations, it is expected values of functions of the payoffs that have to be compared for the resident strategy to be SLS, either the expected values of the logarithm of the payoffs against the resident strategy or, in case of equality of these payoffs, the expected values of the ratio of the payoffs against the mutant strategy over the common payoff against the resident strategy. Assuming small enough perturbations, these conditions can be expressed in terms of means, variances and covariances of the payoffs.

Under conditions on the random payoffs for a constant interior equilibrium to exist, we have found a condition for this equilibrium to be SLS. We have shown that this equilibrium and both fixation states can be simultaneously SLS. This situation distinguishes game dynamics in a randomly fluctuating environment from game dynamics in a constant environment since, with constant payoffs, an interior equilibrium can be locally stable only if both fixation states are locally unstable (Lessard, 1984).

In a constant environment, an evolutionarily stable strategy (ESS) and a convergence stable strategy (CSS) with respect to mixed strategies on two pure strategies correspond to a locally stable equilibrium with respect to the dynamics involving the two pure strategies. Both evolutionary concepts have been extended to take into account random perturbations on payoffs by using SLS and SLU conditions. We have shown that the condition for a pure strategy to be stochastically evolutionarily stable (SES) and stochastically convergence stable (SCS) is more stringent than in a constant environment, while the condition for a constant mixed strategy to be SES is less stringent than the condition to be SCS which is less stringent than the condition in a constant environment.

New phenomenons arise in game dynamics in a stochastic environment, and these make it not only more complex but also more interesting.

1.6 Appendix

1.6.1 A. Theorem 1

Theorem 1.1 *The fixation state* $\hat{x} = 0$ *of the recurrence equation (Eq. (1.2.5)) with the payoff matrix (Eq. (1.2.1)) is stochastically locally stable if*

$$\mathbb{E}\left(\log\left(\frac{d_t}{b_t}\right)\right) \;=\; \mathbb{E}\left(\log d_t\right) - \mathbb{E}\left(\log b_t\right) \;>\; 0\ , \tag{1.6.1}$$

and stochastically locally unstable if the inequality is reversed.

We follow Karlin and Liberman (1975) but with non-symmetric fitness parameters given by the entries of the payoff matrix Eq. (1.2.1). It is easy to check that Eq. (1.2.7) can be written in the form

$$\frac{u_{t+1}}{u_t} \quad = \quad \frac{b_t}{d_t}\left(1+\frac{u_t(a_t d_t - b_t c_t)}{u_t b_t c_t + b_t d_t}\right) \ , \tag{1.6.2}$$

from which

$$\frac{1}{n}\left(\log u_n - \log u_0\right) \quad = \quad \frac{1}{n}\sum_{t=0}^{n-1}\log\left(\frac{b_t}{d_t}\right) + \frac{1}{n}\sum_{t=0}^{n-1}\log\left(1+\frac{u_t(a_t d_t - b_t c_t)}{u_t b_t c_t + b_t d_t}\right) \ , \tag{1.6.3}$$

for $n \geq 1$. Let

$$\mu \quad = \quad \mathbb{E}\left(\log\left(\frac{b_t}{d_t}\right)\right) \quad = \quad \mathbb{E}\left(\log b_t\right) - \mathbb{E}\left(\log d_t\right) \ , \tag{1.6.4}$$

and define

$$E \quad = \quad \left\{\frac{1}{n}\sum_{t=0}^{n-1}\log\left(\frac{b_t}{d_t}\right) \to \mu\right\} \ . \tag{1.6.5}$$

The strong law of large numbers garantees that $\mathbb{P}(E) = 1$. If $u_t \to 0$, then

$$\log\left(1+\frac{u_t(a_t d_t - b_t c_t)}{u_t b_t c_t + b_t d_t}\right) \quad \to \quad 0 \ , \tag{1.6.6}$$

since a_t, b_t, c_t, d_t are assumed to be uniformly bounded below and above by positive constants. Under these conditions, Eq. (1.6.3) implies that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{t=0}^{n-1}\log\left(\frac{b_t}{d_t}\right) \quad \leq \quad 0 \tag{1.6.7}$$

if this limit exists. This is not possible in the set E if $\mu > 0$. In this case, we conclude that

$$\mathbb{P}(u_t \to 0) \quad \leq \quad \mathbb{P}(E^C) \quad = \quad 0 \ . \tag{1.6.8}$$

This means that $\hat{u} = 0$ is stochastically locally unstable if $\mu > 0$.

Now consider the case where $\mu < 0$. By the strong law of large numbers and Egorov's theorem, for any $\epsilon > 0$, there exists an integer $N \geq 1$ such that the probability of the event

$$F = \left\{ \frac{1}{n} \sum_{t=0}^{n-1} \log\left(\frac{b_t}{d_t}\right) < \frac{\mu}{2}, \ \forall\, n \geq N \right\} \tag{1.6.9}$$

satisfies

$$\mathbb{P}(F) \geq 1 - \epsilon\ . \tag{1.6.10}$$

On the other hand, using the assumption that $A \leq a_t, b_t, c_t, d_t \leq B$ for some constants $A, B > 0$, there exists $\delta > 0$ such that

$$\log\left(1 + \frac{u_t(a_t d_t - b_t c_t)}{u_t b_t c_t + b_t d_t}\right) < -\frac{\mu}{4} \tag{1.6.11}$$

as soon as $u_t < \delta$. Moreover, Eq. (1.2.7) leads to

$$u_{t+1} \leq u_t \left(\frac{u_t B + B}{u_t A + A}\right) = u_t \left(\frac{B}{A}\right) \leq u_0 \left(\frac{B}{A}\right)^{t+1}, \tag{1.6.12}$$

for $t \geq 0$. Therefore, there exists $0 < \delta_0 < \delta$ such that $u_t < \delta$ for $t = 0, 1, \ldots, N-1$ as soon as $u_0 < \delta_0$. As a consequence, Eq. (1.6.3) for $n = N$ and Eq. (1.6.11) yield

$$\frac{1}{N}(\log u_N - \log u_0) < \frac{\mu}{2} - \frac{\mu}{4} = \frac{\mu}{4} < 0 \tag{1.6.13}$$

in the set F as soon as $u_0 < \delta_0$, which implies that

$$u_N < u_0 < \delta, \tag{1.6.14}$$

and by recurrence that $u_n < \delta$ for all $n \geq N$.

It remains to show that $u_n \to 0$ in F if $u_0 < \delta_0$ as claimed in Karlin and Liberman (1975), since then

$$P(u_n \to 0) \geq P(F) \geq 1 - \epsilon\ . \tag{1.6.15}$$

It suffices to notice that Eq. (1.6.3) for all $n \geq N$ under the above conditions gives

$$\frac{1}{n}(\log u_n - \log u_0) < \frac{\mu}{4} < 0\ , \tag{1.6.16}$$

from which

$$\log u_n < \log u_0 + \frac{n\mu}{4} \to -\infty. \tag{1.6.17}$$

This means that $u_n \to 0$, which completes the proof.

1.6.2 B. Theorem 2

Theorem 1.2 *The fixation state* $\hat{x} = 0$ *of the recurrence equation (Eq. (1.2.5)) with the payoff matrix (Eq. (1.2.1)) in the case where* $b_t = d_t$ *for all* $t \geq 0$ *is stochastically locally stable if*

$$\mathbb{E}\left(\frac{c_t}{d_t} - \frac{a_t}{d_t}\right) \;=\; \mathbb{E}\left(\frac{c_t}{d_t}\right) - \mathbb{E}\left(\frac{a_t}{d_t}\right) \;>\; 0 \;, \tag{1.6.18}$$

and stochastically locally unstable if the inequality is reversed.

Assuming $b_t = d_t$, the recurrence equation Eq. (1.2.7) with the change of variables $v_t = 1/u_t$ becomes

$$v_{t+1} \;=\; v_t\left(\frac{c_t + d_t v_t}{a_t + d_t v_t}\right) \;, \tag{1.6.19}$$

from which

$$v_{t+1} - v_t \;=\; \left(\frac{c_t}{d_t} - \frac{a_t}{d_t}\right) - \frac{\frac{a_t}{d_t}\left(1 - \frac{a_t}{c_t}\right)}{\frac{a_t}{c_t} + \frac{d_t}{c_t} v_t} \tag{1.6.20}$$

for $t \geq 0$, and therefore

$$\frac{1}{n}\left(v_n - v_0\right) \;=\; \frac{1}{n}\sum_{t=0}^{n-1}\left(\frac{c_t}{d_t} - \frac{a_t}{d_t}\right) - \frac{1}{n}\sum_{t=0}^{n-1}\frac{\frac{a_t}{d_t}\left(1 - \frac{a_t}{c_t}\right)}{\frac{a_t}{c_t} + \frac{d_t}{c_t} v_t} \tag{1.6.21}$$

for $n \geq 1$. Defining

$$E \;=\; \left\{\frac{1}{n}\sum_{t=0}^{n-1}\left(\frac{c_t}{d_t} - \frac{a_t}{d_t}\right) \to \mu\right\} \;, \tag{1.6.22}$$

where

$$\mu \;=\; \mathbb{E}\left(\frac{c_t}{d_t} - \frac{a_t}{d_t}\right) \;=\; \mathbb{E}\left(\frac{c_t}{d_t}\right) - \mathbb{E}\left(\frac{a_t}{d_t}\right) \;, \tag{1.6.23}$$

we conclude as in the proof of Theorem 1 that

$$\mathbb{P}(v_t \to +\infty) \;\leq\; \mathbb{P}(E^C) \;=\; 0 \tag{1.6.24}$$

if $\mu < 0$. On the other hand, if $\mu > 0$, then there exist an integer $N \geq 1$ and a real number $\Delta > 0$ such that

$$F \;=\; \left\{ \frac{1}{n} \sum_{t=0}^{n-1} \left(\frac{c_t}{d_t} - \frac{a_t}{d_t} \right) > \frac{\mu}{2},\ \forall\, n \geq N \right\} \tag{1.6.25}$$

satisfies $\mathbb{P}(F) \geq 1 - \epsilon$ for any given $\epsilon > 0$, and

$$-\frac{\frac{a_t}{d_t}\left(1 - \frac{a_t}{c_t}\right)}{\frac{a_t}{c_t} + \frac{d_t}{c_t} v_t} \;>\; -\frac{\mu}{4} \tag{1.6.26}$$

as soon as $v_t > \Delta$, which is the case for $t = 0, 1, \ldots, N-1$ as soon as $v_0 > \Delta_0$ for some $\Delta_0 > \Delta$ since

$$v_{t+1} \;\geq\; v_t \left(\frac{A + Av_t}{B + Bv_t} \right) \;=\; v_t \left(\frac{A}{B} \right) \;\geq\; v_0 \left(\frac{A}{B} \right)^{t+1} \;\geq\; v_0 \left(\frac{A}{B} \right)^{N} . \tag{1.6.27}$$

Then, as in the proof of Theorem 1, it can be shown that we have $v_n > \Delta$ for all $n \geq N$ and $v_n \to +\infty$ in F as soon as $v_0 > \Delta_0$, from which

$$P(v_n \to +\infty) \;\geq\; P(F) \geq 1 - \epsilon \tag{1.6.28}$$

as soon as $v_0 > \Delta_0$.

1.6.3 C. Theorem 3

Theorem 1.3 *A constant equilibrium* $\hat{x} = \hat{u}/(1+\hat{u})$ *of the recurrence equation (Eq. (1.2.5)) with* $\hat{u} > 0$ *and the payoff matrix (Eq. (1.3.17)) is stochastically locally stable if*

$$\mathbb{E}\left(\log \left(\frac{\hat{u}c_t + d_t}{\hat{u}a_t + d_t} \right) \right) \;=\; \mathbb{E}\left(\log \left(1 - \frac{\hat{x} z_t}{\hat{x} a_t + (1-\hat{x}) d_t} \right) \right) \;>\; 0\ , \tag{1.6.29}$$

and stochastically locally unstable if the inequality is reversed.

With the payoffs given by the entries of the game matrix (Eq. (1.3.17)), the recurrence equation (Eq. (1.2.5)) can be written in the form

$$u_{t+1} \;=\; u_t \left(\frac{u_t c_t + u_t z_t + b_t}{u_t c_t + \hat{u} z_t + b_t} \right)\ , \tag{1.6.30}$$

from which it is easy to get

$$\begin{aligned} u_{t+1}-\hat{u} &= (u_t-\hat{u})\left(\frac{u_tc_t+u_tz_t+\hat{u}z_t+b_t}{u_tc_t+\hat{u}z_t+b_t}\right) \\ &= (u_t-\hat{u})\left(\frac{u_ta_t+d_t}{u_tc_t+d_t}\right) \ . \end{aligned} \tag{1.6.31}$$

In particular, this ensures that $u_{t+1}-\hat{u}>0$ if $u_t-\hat{u}>0$, and $u_{t+1}-\hat{u}<0$ if $u_t-\hat{u}<0$. Moreover, some algebraic manipulations lead to

$$\frac{u_ta_t+d_t}{u_tc_t+d_t} = \left(\frac{\hat{u}a_t+d_t}{\hat{u}c_t+d_t}\right)\left(1-\frac{d_tz_t(u_t-\hat{u})}{(\hat{u}a_t+d_t)(\hat{u}c_t+d_t+(u_t-\hat{u})c_t)}\right) \ . \tag{1.6.32}$$

In order to conclude, it suffices to proceed as in the proof of Theorem 1 and to note that

$$\begin{aligned} \log\left(\frac{\hat{u}a_t+d_t}{\hat{u}c_t+d_t}\right) &= -\log\left(\frac{\hat{u}c_t+d_t}{\hat{u}a_t+d_t}\right) \\ &= -\log\left(1-\frac{\hat{x}z_t}{\hat{x}a_t+(1-\hat{x})d_t}\right) \ . \end{aligned} \tag{1.6.33}$$

1.6.4 D. Proof of Result 1

For $\hat{\mathbf{x}}=(0,1)$, $\mathbf{x}=(x,1-x)$ and $\mathbf{A}(t)=\bar{\mathbf{A}}+\mathbf{B}(t)$ as in Eq. (1.4.1), we find

$$\begin{aligned} \bar{d} &= \mathbb{E}\left(\hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}}\right) = \bar{a}_{22} \ , \\ \bar{b} &= \mathbb{E}\left(\mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}}\right) = \bar{a}_{22}+x(\bar{a}_{12}-\bar{a}_{22}) \ , \\ \sigma_d^2 &= \mathbb{E}\left((\hat{\mathbf{x}}\mathbf{B}(t)\hat{\mathbf{x}})^2\right) = \sigma_{22}^2 \ , \\ \sigma_b^2 &= \mathbb{E}\left((\mathbf{x}\mathbf{B}(t)\hat{\mathbf{x}})^2\right) = (1-x)^2\sigma_{22}^2+x^2\sigma_{12}^2 \ . \end{aligned} \tag{1.6.34}$$

Condition Eq. (1.3.5) for $\hat{\mathbf{x}}$-fixation to be SLS against $\mathbf{x}$ if the variances are small enough becomes

$$\log\left(1+x\left(\frac{\bar{a}_{12}-\bar{a}_{22}}{\bar{a}_{22}}\right)\right) < \frac{1}{2}\left(\frac{(1-x)^2\sigma_{22}^2+x^2\sigma_{12}^2}{(\bar{a}_{22}+x(\bar{a}_{12}-\bar{a}_{22}))^2}-\frac{\sigma_{22}^2}{\bar{a}_{22}^2}\right) \ . \tag{1.6.35}$$

This condition reduces to

$$x\left(\frac{\bar{a}_{12}-\bar{a}_{22}}{\bar{a}_{22}}\right) < -x\left(\frac{\bar{a}_{12}\sigma_{22}^2}{\bar{a}_{22}^3}\right) \tag{1.6.36}$$

for $x > 0$ small enough. This condition is equivalent to

$$\overline{a}_{22}^2\,(\overline{a}_{12} - \overline{a}_{22}) \quad > \quad \overline{a}_{12}\sigma_{22}^2, \tag{1.6.37}$$

which is the same as condition Eq. (1.4.3). The reversed inequality ensures that $\hat{\mathbf{x}}$-fixation is SLU.

Next, we study stochastic convergence stability. Consider a strategy $\tilde{\mathbf{x}} = (\tilde{x}, 1-\tilde{x})$ with $\tilde{x} > 0$ near the pure strategy $\hat{\mathbf{x}} = (\hat{x}, 1-\hat{x}) = (0,1)$. Given another strategy $\mathbf{x} = (x, 1-x)$, the payoff matrix for $\mathbf{x}$ and $\tilde{\mathbf{x}}$ in this order is given by

$$\begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} \quad = \quad \begin{pmatrix} \mathbf{x}\mathbf{A}(t)\mathbf{x} & \mathbf{x}\mathbf{A}(t)\tilde{\mathbf{x}} \\ \tilde{\mathbf{x}}\mathbf{A}(t)\mathbf{x} & \tilde{\mathbf{x}}\mathbf{A}(t)\tilde{\mathbf{x}} \end{pmatrix} \ , \tag{1.6.38}$$

where $\mathbf{A}(t) = \bar{\mathbf{A}} + \mathbf{B}(t)$ as in Eq. (1.4.1). We find

$$\begin{aligned} \bar{d} \ &= \ \mathbb{E}\left(\tilde{\mathbf{x}}\mathbf{A}(t)\tilde{\mathbf{x}}\right) \ = \ \tilde{x}^2\bar{a}_{11} + \tilde{x}(1-\tilde{x})(\bar{a}_{12} + \bar{a}_{21}) + (1-\tilde{x})^2\bar{a}_{22} \ , \\ \bar{b} \ &= \ \mathbb{E}\left(\mathbf{x}\mathbf{A}(t)\tilde{\mathbf{x}}\right) \ = \ x\tilde{x}\bar{a}_{11} + x(1-\tilde{x})\bar{a}_{12} + (1-x)\tilde{x}\bar{a}_{21} + (1-x)(1-\tilde{x})\bar{a}_{22} \ , \end{aligned} \tag{1.6.39}$$

and

$$\begin{aligned} \sigma_d^2 \ &= \ \mathbb{E}\left((\tilde{\mathbf{x}}\mathbf{B}(t)\tilde{\mathbf{x}})^2\right) \\ &= \ \tilde{x}^4\sigma_{11}^2 + \tilde{x}^2(1-\tilde{x})^2(\sigma_{12}^2 + \sigma_{21}^2) + (1-\tilde{x})^4\sigma_{22}^2 \ , \\ \sigma_b^2 \ &= \ \mathbb{E}\left((\mathbf{x}\mathbf{B}(t)\tilde{\mathbf{x}})^2\right) \\ &= \ x^2\tilde{x}^2\sigma_{11}^2 + x^2(1-\tilde{x})^2\sigma_{12}^2 + (1-x)^2\tilde{x}^2\sigma_{21}^2 + (1-x)^2(1-\tilde{x})^2\sigma_{22}^2 \ . \end{aligned} \tag{1.6.40}$$

Defining $\Delta x = x - \tilde{x}$, the above expressions lead to

$$\begin{aligned} \bar{b} - \bar{d} \ &= \ -g_1(\tilde{x})\Delta x \ , \\ \sigma_b^2 - \sigma_d^2 \ &= \ -2h_1(\tilde{x})\Delta x + o(\Delta x) \ , \end{aligned} \tag{1.6.41}$$

where

$$g_1(\tilde{x}) \ = \ (\bar{a}_{22} - \bar{a}_{12}) + (\bar{a}_{21} - \bar{a}_{11} + \bar{a}_{12} - \bar{a}_{22})\tilde{x} \ , \tag{1.6.42}$$

$$h_1(\tilde{x}) \ = \ -\tilde{x}^3\sigma_{11}^2 - \tilde{x}(1-\tilde{x})\sigma_{12}^2 + \tilde{x}^2(1-\tilde{x})\sigma_{21}^2 + (1-\tilde{x})^3\sigma_{22}^2 \ . \tag{1.6.43}$$

With small enough variances, strategy $\tilde{\mathbf{x}}$ is SLU against strategy $\mathbf{x}$ if

$$\log\left(\frac{\bar{b}}{\bar{d}}\right) \quad > \quad \frac{1}{2}\left(\frac{\sigma_b^2}{\bar{b}^2} - \frac{\sigma_d^2}{\bar{d}^2}\right) \ , \tag{1.6.44}$$

where

$$\begin{aligned} \log\left(\frac{\bar{b}}{\bar{d}}\right) &= \log\left(1 + \frac{\bar{b}-\bar{d}}{\bar{d}}\right) \\ &= \log\left(1 - \frac{g_1(\tilde{x})}{\bar{d}}\Delta x\right) \\ &= -\frac{g_1(\tilde{x})}{\bar{d}}\Delta x + o(\Delta x) \ , \end{aligned} \tag{1.6.45}$$

while

$$\begin{aligned} \frac{1}{2}\left(\frac{\sigma_b^2}{\bar{b}^2} - \frac{\sigma_d^2}{\bar{d}^2}\right) &= \frac{\sigma_b^2}{2\bar{d}^2}\left(\frac{1}{1+(\bar{b}-\bar{d})/\bar{d}}\right)^2 - \frac{\sigma_d^2}{2\bar{d}^2} \\ &= \frac{\sigma_b^2}{2\bar{d}^2}\left(1 + \frac{g_1(\tilde{x})}{\bar{d}}\Delta x\right)^2 - \frac{\sigma_d^2}{2\bar{d}^2} + o(\Delta x) \\ &= \frac{\sigma_b^2 - \sigma_d^2}{2\bar{d}^2} + \frac{\sigma_b^2 g_1(\tilde{x})}{\bar{d}^3}\Delta x + o(\Delta x) \\ &= -\frac{h_1(\tilde{x})}{\bar{d}^2}\Delta x + \frac{\sigma_d^2 g_1(\tilde{x})}{\bar{d}^3}\Delta x + o(\Delta x) \ . \end{aligned} \tag{1.6.46}$$

If $\Delta x < 0$ and $|\Delta x|$ small enough, then condition Eq. (1.6.44) is equivalent to

$$(\bar{d}^2 + \sigma_d^2)g_1(\tilde{x}) > \bar{d}h_1(\tilde{x}) \ . \tag{1.6.47}$$

For $\tilde{\mathbf{x}}$ close enough to $\hat{\mathbf{x}}$, this condition reduces to

$$(\bar{a}_{22}^2 + \sigma_{22}^2)(\bar{a}_{22} - \bar{a}_{12}) > \sigma_{22}^2 \bar{a}_{22} \ , \tag{1.6.48}$$

which is equivalent to Eq. (1.4.3). On the contrary, if $\Delta x > 0$, then strategy $\tilde{\mathbf{x}}$ is SLS against strategy $\mathbf{x}$.

1.6.5 E. Proof of Result 2

Dropping the mention of the time step to simplify the notation, let $\mathbf{A} = (a_{ij})$ be a stochastic game matrix with $a_{11} = a_{21} + z$ and $a_{22} = a_{12} + \hat{u}z$ where $\hat{u} = \hat{x}/(1-\hat{x}) > 0$.

The payoff of the mixed strategy $\mathbf{x} = (x, 1-x)$ against $\hat{\mathbf{x}} = (\hat{x}, 1-\hat{x})$ is then given by

$$b \;=\; \mathbf{x}\mathbf{A}\hat{\mathbf{x}} = a_{21}\hat{x} + z\hat{x} + a_{12}(1-\hat{x}) \;, \tag{1.6.49}$$

which does not depend on $\mathbf{x}$. Therefore, it is the same as the payoff of $\hat{\mathbf{x}}$ against itself, that is

$$d \;=\; \hat{\mathbf{x}}\mathbf{A}\hat{\mathbf{x}} \;=\; b \;. \tag{1.6.50}$$

On the other hand, the payoff of $\hat{\mathbf{x}} = (\hat{x}, 1-\hat{x})$ against $\mathbf{x} = (x, 1-x)$ is

$$c \;=\; \hat{\mathbf{x}}\mathbf{A}\mathbf{x} \;=\; \hat{\mathbf{x}}\mathbf{A}(\mathbf{x}-\hat{\mathbf{x}}) + \hat{\mathbf{x}}\mathbf{A}\hat{\mathbf{x}} \;=\; \hat{\mathbf{x}}\mathbf{A}\boldsymbol{\delta} + d \;, \tag{1.6.51}$$

where $\boldsymbol{\delta} = (\delta, -\delta) = (x-\hat{x}, \hat{x}-x) = \mathbf{x}-\hat{\mathbf{x}}$, while the payoff of $\mathbf{x} = (x, 1-x)$ against itself is

$$\begin{aligned} a \;&=\; \mathbf{x}\mathbf{A}\mathbf{x} \;=\; (\mathbf{x}-\hat{\mathbf{x}})\mathbf{A}(\mathbf{x}-\hat{\mathbf{x}}) + \hat{\mathbf{x}}\mathbf{A}(\mathbf{x}-\hat{\mathbf{x}}) + \mathbf{x}\mathbf{A}\hat{\mathbf{x}} \\ &=\; \boldsymbol{\delta}\mathbf{A}\boldsymbol{\delta} + \hat{\mathbf{x}}\mathbf{A}\boldsymbol{\delta} + d \;=\; \boldsymbol{\delta}\mathbf{A}\boldsymbol{\delta} + c \;. \end{aligned} \tag{1.6.52}$$

Note that

$$\boldsymbol{\delta}\mathbf{A}\boldsymbol{\delta} \;=\; \delta^2(a_{11} - a_{12} - a_{21} + a_{22}) \;=\; \delta^2(1+\hat{u})z \;. \tag{1.6.53}$$

Therefore,

$$\mathbb{E}\left(\frac{a}{d} - \frac{c}{d}\right) \;=\; \mathbb{E}\left(\frac{\boldsymbol{\delta}\mathbf{A}\boldsymbol{\delta}}{d}\right) \;=\; \delta^2(1+\hat{u})\mathbb{E}\left(\frac{z}{d}\right) \;. \tag{1.6.54}$$

According to Theorem 2, $\hat{\mathbf{x}}$-fixation is SLS if

$$\mathbb{E}\left(\frac{z}{d}\right) \;<\; 0 \;, \tag{1.6.55}$$

and SLU if this inequality is reversed.

Let us write

$$\begin{aligned} a_{12} \;&=\; \bar{a}_{12} + \eta_{12} \;, \\ a_{21} \;&=\; \bar{a}_{21} + \eta_{21} \;, \\ z \;&=\; \bar{z} + \zeta \;, \end{aligned} \tag{1.6.56}$$

where η_{21}, η_{12} and ζ are independent random variables of mean 0 and variances σ_{21}^2, σ_{12}^2 and σ_z^2, respectively, while all higher-order centered moments are $o(\sigma^2)$ where $\sigma^2 = \max\{\sigma_{21}^2, \sigma_{12}^2, \sigma_z^2\}$. Then we find

$$\mathbb{E}\left(\frac{z}{d}\right) = \frac{1}{\bar{d}}\left(\bar{z}\bar{d}^2 + \bar{z}\sigma_{21}^2\hat{x}^2 + \bar{z}\sigma_{12}^2(1-\hat{x})^2 - \bar{d}\sigma_z^2\hat{x}\right) + o(\sigma^2) \ , \tag{1.6.57}$$

where

$$\bar{d} = \bar{z}\hat{x} + \bar{a}_{21}\hat{x} + \bar{a}_{12}(1-\hat{x}) \ . \tag{1.6.58}$$

In the case where σ^2 is small enough, we conclude that $\hat{\mathbf{x}}$-fixation is SLS if

$$\bar{z}\bar{d}^2 + \bar{z}\sigma_{21}^2\hat{x}^2 + \bar{z}\sigma_{12}^2(1-\hat{x})^2 - \bar{d}\sigma_z^2\hat{x} < 0 \ , \tag{1.6.59}$$

and SLU if this inequality is reversed. This gives the condition for $\hat{\mathbf{x}}$ to be a stochastically evolutionarily stable strategy (SESS).

For stochastic convergence stability (SCS) of $\hat{\mathbf{x}}$, we consider a strategy $\tilde{\mathbf{x}} = (\tilde{x}, 1-\tilde{x})$ near $\hat{\mathbf{x}} = (\hat{x}, 1-\hat{x})$. Given another strategy $\mathbf{x} = (x, 1-x)$, the payoff matrix for $\mathbf{x}$ and $\tilde{\mathbf{x}}$ in this order is, again, given by Eq. (1.6.38). Therefore, the means of the variables b and d, $\bar{b}$ and $\bar{d}$, have the same expressions as in Eq. (1.6.39), but with $\bar{a}_{11} = \bar{a}_{21} + \bar{z}$ and $\bar{a}_{22} = \bar{a}_{12} + \hat{u}\bar{z}$. Defining $\Delta x = x - \tilde{x}$ and $\Delta\hat{x} = \hat{x} - \tilde{x}$, the difference of the means is found to be

$$\bar{b} - \bar{d} = -g_2(\tilde{x})\Delta x \ , \tag{1.6.60}$$

where

$$g_2(\tilde{x}) = \bar{z}\big(\hat{u} - (1+\hat{u})\tilde{x}\big) = (1+\hat{u})\Delta\hat{x} \ . \tag{1.6.61}$$

On the other hand, since

$$\begin{aligned} d &= \tilde{x}^2(a_{21}+z) + \tilde{x}(1-\tilde{x})(a_{12}+a_{21}) + (1-\tilde{x})^2(a_{12}+\hat{u}z) \\ &= (1-\tilde{x})a_{12} + \tilde{x}a_{21} + \big(\tilde{x}^2 + \hat{u}(1-\tilde{x})^2\big)z \ , \\ b &= x\tilde{x}(a_{21}+z) + x(1-\tilde{x})a_{12} + \tilde{x}(1-x)a_{21} + (1-x)(1-\tilde{x})(a_{12}+\hat{u}z) \\ &= (1-\tilde{x})a_{12} + \tilde{x}a_{21} + \big(x\tilde{x} + \hat{u}(1-x)(1-\tilde{x})\big)z \ , \end{aligned} \tag{1.6.62}$$

where a_{12}, a_{21} and z are independent random variables, the variances of b and d are given by

$$\begin{aligned} \sigma_d^2 &= (1-\tilde{x})^2\sigma_{12}^2 + \tilde{x}^2\sigma_{21}^2 + (\tilde{x}^2 + \hat{u}(1-\tilde{x})^2)^2\sigma_z^2 \ , \\ \sigma_b^2 &= (1-\tilde{x})^2\sigma_{12}^2 + \tilde{x}^2\sigma_{21}^2 + (x\tilde{x} + \hat{u}(1-x)(1-\tilde{x}))^2\sigma_z^2 \ . \end{aligned} \tag{1.6.63}$$

Writing $x = \tilde{x} + \Delta x$ and $\tilde{x} = \hat{x} - \Delta\hat{x}$, the difference of the variances is found to be

$$\sigma_b^2 - \sigma_d^2 = -2h_2(\tilde{x})\Delta x + o(\Delta x) \ , \tag{1.6.64}$$

where

$$\begin{aligned} h_2(\tilde{x}) &= \sigma_z^2\big(\tilde{x}^2 + \hat{u}(1-\tilde{x})^2\big)\big(\hat{u}(1-\tilde{x}) - \tilde{x}\big) \\ &= \sigma_z^2\big(\hat{x}^2 + \hat{u}(1-\hat{x})^2\big)\Delta\hat{x} + o(\Delta\hat{x}) \\ &= \sigma_z^2(1+\hat{u})\hat{x}\Delta\hat{x} + o(\Delta\hat{x}) \ . \end{aligned} \tag{1.6.65}$$

Analogously to the conclusion drawn in the proof of Result 1, if Δx in absolute value and the variances are small enough, then strategy $\tilde{\mathbf{x}}$ is SLU against strategy $\mathbf{x}$ if

$$(\bar{d}^2 + \sigma_d^2)g_2(\tilde{x})\Delta x < \bar{d}h_2(\tilde{x})\Delta x \ . \tag{1.6.66}$$

where

$$\sigma_d^2 = (1-\hat{x})^2\sigma_{12}^2 + \hat{x}^2\sigma_{21}^2 + \hat{x}^2\sigma_z^2 + O(\Delta\hat{x}) \ . \tag{1.6.67}$$

If $\Delta\hat{x}$ is small enough and such that $(\Delta x)(\Delta\hat{x}) > 0$, which means that $\mathrm{sgn}(x - \tilde{x}) = \mathrm{sgn}(\hat{x} - \tilde{x})$, then Eq. (1.6.66) reduces to

$$\bar{z}\bar{d}^2 + \bar{z}\sigma_{21}^2\hat{x}^2 + \bar{z}\sigma_{12}^2(1-\hat{x})^2 + \bar{z}\hat{x}^2\sigma_z^2 - \bar{d}\sigma_z^2\hat{x} < 0 \ . \tag{1.6.68}$$

If either $\Delta\hat{x}$ or Δx changes sign, then the reverse inequality is obtained, which means that strategy $\tilde{\mathbf{x}}$ is SLS against strategy $\mathbf{x}$.

Acknowledgements

In this study, X-D.Z. and Y.T. were supported by National Natural Science Foundation of China (11401562 and 11471311, respectively); S.L. was supported in part by Natural Sciences and Engineering Research Council of Canada, Grant no. 8833.

Bibliography

[1] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, UK, 1982).

[2] J. Maynard Smith, G.R. Price, The logic of animal conflicts. *Nature* **246**, 15-18 (1973).

[3] J. Maynard Smith, The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology* **47**, 209-221 (1974).

[4] P.D. Taylor, L. Jonker, Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* **40**, 145-156 (1978).

[5] J. Hofbauer, K. Sigmund, *The Theory of Evolution and Dynamical Systems.* (Cambridge University Press, Cambridge, UK, 1998).

[6] M.A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646-650 (2004).

[7] M. Schaffer, Evolutionarily stable strategies for a finite population and a variable contest size. *Journal of Theoretical Biology* **132**, 469-478 (1988).

[8] S. Lessard, V. Ladret, The probability of fixation of a single mutant in an exchangeable selection model. *Journal of Mathematical Biology* **54**, 721-744 (2007).

[9] T. Antal, M.A. Nowak, A. Traulsen, Strategy abundance in 2×2 games for arbitrary mutation rates. *Journal of Theoretical Biology* **257**, 340-344 (2009).

[10] D. Foster and P. Young, Stochastic evolutionary game dynamics. *Theoretical Population Biology* **38**, 219-232 (1990).

[11] P. Turchin, L. Oksanen, P. Ekerholm, T. Oksanen, H. Henttonen, Are lemmings prey or predators? *Nature* **405**, 562 (2000).

[12] O.N. Bjornstad, B.T. Grenfell, Noisy clockwork: Time series analysis of population fluctuations in animals *Science* **293**, 638-642 (2001).

[13] E.M. Ozbudak, M. Thattai, I. Kurtser, A.D. Grossman, A. van Oudenaarden, *Nature Genetics* **31**, 69-73 (2002).

[14] M.B. Elowitz, A.J. Levine, E.D. Siggia, P.S. Swain, stochastic Gene expression in a single cell. *Science* **297**, 1183-1186 (2002).

[15] W.J. Blake, M. Kaern, C.R. Cantor, J.J. Collins, Noise in eukaryotic gene expression. *Nature* **422**, 633-637 (2003).

[16] J. Paulsson, Summing up the noise in gene networks. *Nature* **427**, 415-418 (2004).

[17] P. Romanczuk, I.D. Couzin, L. Schimansky-Geier, Collective motion due to individual escape and pursuit response. *Physical Review Letters* **102**, 010602 (2009).

[18] S. Ciuchi, F. de Pasquale, B. Spagnolo, Nonlinear relaxation in the presence of an absorbing barrier. *Physical Review E* **47**, 3915-3926 (1993).

[19] B. Spagnolo, A. Fiasconaro, D. Valenti, Noise induced phenomena in Lotka-Volterra systems. *Fluctuation and Noise Letters* **3**, L177-L185 (2003).

[20] B. Spagnolo, D. Valenti and A. Fiasconaro, Noise in ecosystem: a short review. *Mathematical Biosciences and Engineering* **1**, 185-211 (2004).

[21] O. Chichigina, D. Valenti, B. Spagnolo, A simple noise model with memory for biological systems. *Fluctuation and Noise Letters* **5**, L243-L250 (2005).

[22] A. La Cognata, D. Valenti, A.A. Dubkov, B. Spagnolo, Dynamics of two competing species in the presence of Levy noise sources. *Physical Review E* **82**, 011121 (2010).

[23] O.A. Chichigina, A.A. Dubkov, D. Valenti, B. Spagnolo, Stability in a system subject to noise with regulated periodicity. *Physical Review E* **84**, 021134 (2011).

[24] I. Eshel and U. Motro, Kin selection and strong evolutionary stability of mutual help. *Theoretical Population Biology* **19**, 420-433 (1981).

[25] I. Eshel, Evolutionary and continuous stability. *Journal of Theoretical Biology* **103**, 99-111 (1983).

[26] I. Eshel, On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution. *Journal of Mathematical Biology* **34**, 485-510 (1996).

[27] F.B. Christiansen, On conditions for evolutionary stability for a continuously varying character. *American Naturalist* **138**, 37-50 (1991).

[28] S. Karlin, U. Liberman, Random temporal variation in selection intensities: case of large population size. *Theoretical Population Biology* **6**, 355-382 (1974).

[29] S. Karlin, U. Liberman, Random temporal variation in selection intensities: One-locus two-allele model. *Journal of Mathematical Biology* **2**, 1-17 (1975).

[30] W.J. Ewens, *Mathematical Population Genetics: I Theoretical Introduciton* (Springer, New York, USA, 2004).

[31] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[32] M.A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).

[33] S. Lessard, Evolutionary dynamics in frequency-dependent two-phenotype models. *Theoretical Population Biology* **25**, 210-234 (1984).

Chapter 2

Article 2

Environmental noise could promote stochastic local stability of behavioral diversity evolution

Xiu-Deng Zheng[1,*], Cong Li[2,*], Sabin Lessard[2] and Yi Tao[1,3]

[1]Key Laboratory of Animal Ecology and Conservation Biology, Centre for Computational and Evolutionary Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, P.R. China

[2]Department of Mathematics and Statistics, University of Montreal, Montreal, Canada

[3]University of Chinese Academy of Sciences, Beijing, China

[*]These authors contributed equally to this work.

Information

This paper appeared in *Physical Review Letters* **120**, 218101. It was received on 5 December 2017; published on 22 May 2018.

My contribution

I contributed in part to the theoretical framework and revised the paper based on a draft by Yi Tao. I ran the numerical simulations for the examples and the figures.

Abstract

In this letter, we investigate stochastic stability in a two-phenotype evolutionary game model for an infinite, well-mixed population undergoing discrete, nonoverlapping generations. We assume that the fitness of a phenotype is an exponential function of its expected payoff following random pairwise interactions whose outcomes randomly fluctuate with time. We show that the stochastic local stability of a constant interior equilibrium can be promoted by the random environmental noise even if the system may display a complicated nonlinear dynamics. This result provides a new perspective for a better understanding of how environmental fluctuations may contribute to the evolution of behavioral diversity.

2.1 Introduction

Since the concept of evolutionarily stable strategy (ESS) has been introduced in evolutionary game theory (Maynard Smith and Price, 1973), it has been successfully applied to explain the evolution of animal behaviors, in particular, altruistic behaviors (or cooperative behaviors) (Maynard Smith, 1982; Axelrod, 1984; Lessard, 1984; Hofbauer and Sigmund, 1988; Nowak, 2006; Broom and Rychtář, 2013). Studies of evolutionary game dynamics start with the replicator equation (Taylor and Jonker, 1978) and focus to a great extent on the time evolution and long-term maintenance of population states pertaining to behavioral diversity (Lessard, 1984; Hofbauer and Sigmund, 1988; Nowak, 2006; Broom and Rychtář, 2013). In the archetypal framework of matrix games in discrete time, there are two important assumptions that are usually made or implicit : the first one is that the fitness of an individual is a simple linear function of the expected payoff of its phenotype, and the second one is that the payoff matrix in pairwise interactions is a constant matrix (Maynard Smith, 1982; Lessard, 1984; Hofbauer and Sigmund, 1988; Nowak, 2006; Broom and Rychtář, 2013). However, both these two assumptions cannot be considered to be always true, or completely real. In this letter, we address the consequences of relaxing these assumptions by considering: (*i*) a random payoff matrix in pairwise interactions, and (*ii*) a nonlinear, actually exponential, fitness function with respect to the expected payoff.

It may be useful to recall the origins of the replicator equation in evolutionary game theory (Taylor and Jonker, 1978; Zeeman, 1980; Maynard Smith, 1982; Hofbauer and Sigmund, 1988). The equation in continuous time was obtained by assuming that the payoff is the current growth rate. If n_i is the current number of i-strategists in a population of total large size $N = \sum_i n_i$, then its time derivative $\dot{n}_i = n_i \pi_i$ where π_i is the payoff to i assumed in general to be frequency-dependent. This leads directly to the replicator equation $\dot{x}_i = x_i(\pi_i - \bar{\pi})$ where $x_i = n_i/N$ is the relative frequency of i and $\bar{\pi}$ the average payoff in the whole population. In discrete time, if $\pi_i(t) \geq -1$ with average value $\bar{\pi}(t)$ is interpreted as the mean number of i-strategists at time step $t+1$ that are produced by each i-strategist at time step t, then the frequency of i satisfies the recurrence equation $x_i(t+1) = \frac{x_i(t)(1+\pi_i(t))}{1+\bar{\pi}(t)}$. Notice that using $c + \pi_i(t)$ (Maynard Smith, 1982; Weibull, 1995) for some positive constant c instead of $1 + \pi_i(t)$ gives the same recurrence equation if all payoffs are multiplied by $1/c$. This multiplicative factor can be interpreted as a strength of selection. Notice also that the aforementioned recurrence equation is a discrete-time approximation of the replicator equation in the case of small payoffs, which means weak selection. To see this, consider payoffs kept constant from time step t to time step $t+1$. This is a reasonable assumption if the time interval is small. Then we have $n_i(t+1) = n_i(t)e^{\pi_i(t)}$, from which we get $x_i(t+1) = \frac{x_i(t)e^{\pi_i(t)}}{\sum_j x_j(t)e^{\pi_j(t)}}$. The approximation $e^{\pi_i(t)} \approx 1 + \pi_i(t)$ in this recurrence equation for $\pi_i(t)$ small gives the previous one. This shows that this recurrence equation is a more general and more precise discrete-time approximation of the replicator equation than the previous one. Moreover, this recurrence equation can be used as an exact discrete-time model for a population with interactions between individuals occurring at the beginning of each time interval (e.g., season) and having effects on growth in number of individuals from the beginning to the end of the time interval. All this is in strong support of fitness in discrete time defined as an exponential function of the payoff, that is, $f_i = e^{\pi_i}$. Notice that such a fitness function is not additive but rather multiplicative. As a consequence, even in the case of an individual payoff that results from random pairwise interactions, it is non linear with respect to the strategy frequencies. As already known, this may lead to dynamical properties very different from, and much more complicated than, those obtained with a linear fitness function even in the case of matrix games with only two phenotypes (Vincent and Fisher, 1988; Blume, 1993; Tao *et al.*, 1997; Szabó and Hauert, 2002; Traulsen *et al.*, 2006; Claussen and Traulsen 2008; Traulsen *et al.*, 2008).

Our objective in this letter is not only to study the effects of a nonlinear fitness function on matrix game dynamics but also the effects of introducing stochastic perturbations of the payoffs. Randomness (or uncertainty) in the environment is one of the main characteristics of nature, and this random noise will generally affect the results of interactions between species and between individuals (May, 1973; Turchin *et al.*, 2000; Lande *et al.*, 2003; Zheng *et al.*, 2017). Therefore, variability in payoffs as measured by their variances and covariances have to be taken into account in order to better understand evolutionary outcomes in natural populations. Below are two examples to show that a random payoff matrix is a reasonable assumption in evolutionary game theory and mathematical ecology.

It is well known that the Lotka-Volterra equation is one of the most important theoretical models in ecology (May, 1973). This equation for the densities of species near equilibrium assumes that the growth rate in continuous-time is density-dependent. In general, the local approximation takes the form $\dot{n}_i = n_i \left(r_i + \sum_{j=1}^{k} a_{ij} n_j \right)$ where (a_{ij}) is an interaction matrix that may depend on environmental carrying capacities. This model is actually equivalent to considering $\tilde{\pi}_i = r_i + N \sum_j a_{ij} x_j$ as the payoff to i, which is not only frequency-dependent but also density-dependent. Then, we would have to consider the time derivative of the population size given by $\dot{N} = N \sum_i x_i \tilde{\pi}_i$ besides the replicator equation in order to describe the whole continuous-time dynamics. Another approach, however, is to introduce an extra species 0 of density $n_0 = 1$ and frequency $x_0 = (N+1)^{-1}$ so that the payoff to i becomes $\tilde{\pi}_i = r_i x_0 + \sum_j a_{ij} x_j$ with respect to the new frequencies and the change of time $\tau = \int_0^t x_0(s) ds$ (Hofbauer and Sigmund, 1988). This leads to a pure frequency-dependent selection model with random pairwise interactions and payoff to i given by a_{ij} when in interaction with j and r_i when in interaction with 0. More importantly, these payoffs are random variables if the carrying capacities are random variables. This is the case with stochastic fluctuations in the environment, which is the rule in nature rather than the exception. In conclusion, extending the Lotka-Volterra equation to take into account a stochastic environment leads to the replicator equation with random payoffs. This is the best example to show that a random payoff matrix is a reasonable assumption. Moreover, this shows that the replicator equation or a discrete-time approximation of it can take into account limiting factors on population size.

Another example is provided by the payoff matrix in the case of repeated rounds of the same matrix game between the same two players chosen at random. This is the case, for instance, with the strategies TFT and AllD in a repeated Prisoner's Dilemma for modeling the evolution of cooperation. If the number of repetitions of the game is a random variable (e.g., a geometric random variable in the case where each round is followed by a next round with some fixed probability (Nowak, 2006; Zheng *et al.*, 2017)), then the payoff matrix is a random matrix. More generally, a random payoff matrix is a reasonable assumption in evolutionary game theory in order to deal with more realistic or more complex situations.

In order to take into account stochastic fluctuations in the surrounding environment, deterministic evolutionary concepts such as evolutionary stability and convergence stability have to be extended. Random payoffs received by randomly pairwise interacting individuals in an infinite population undergoing discrete, nonoverlapping generations were considered till recently in a two-phenotype setting, and the concepts of stochastic evolutionary stability (SES) and stochastic convergence stability were developed (SCS) (Zheng *et al.*, 2017). Although this study shows that a random environmental noise may have an important impact on the stability nature of an equilibrium and, therefore, on the evolution of animal behavior, it still assumes that the fitness of an individual is a linear function of the expected payoff of the exhibited phenotype, which corresponds to an assumption of weak selection. Here we will consider the more general case of an exponential function.

2.2 A two-phenotype model

For simplicity, consider a two-phenotype evolutionary game in an infinite population with discrete, nonoverlapping generations. The two phenotypes (or strategies) are denoted by R_1 and R_2, respectively, and the payoffs in pairwise interactions at time step t (with $t \geq 0$)

are given by the matrix

$$\mathbf{A}(t) \quad = \quad \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix} \quad = \quad \begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} \ , \tag{2.2.1}$$

where $a_{ij}(t)$ is the payoff to strategy R_i against strategy R_j for $i, j = 1, 2$. In general, these payoffs are assumed to be random variables with $\langle a_{ij}(t)\rangle = \bar{a}_{ij}$, $\left\langle (a_{ij}(t) - \bar{a}_{ij})^2 \right\rangle = \sigma_{ij}^2$, and $\langle (a_{ij}(t) - \bar{a}_{ij})\,(a_{kl}(t) - \bar{a}_{kl})\rangle = \sigma_{ij,kl}$ for $i, j, k, l = 1, 2$ with $(i, j) \neq (k, l)$. As for $s \neq t$, the payoffs $a_{ij}(s)$ and $a_{kl}(t)$ are assumed to be independent of each other so that $\langle (a_{ij}(s) - \bar{a}_{ij})\,(a_{kl}(t) - \bar{a}_{kl})\rangle = 0$ for $i, j, k, l = 1, 2$ (Zheng *et al.*, 2017). Besides, a further technical assumption is that there exist real numbers $A, B > 0$ such that $\mathbb{P}\big(A \leq a_{ij}(t) \leq B\big) = 1$ for all $i, j = 1, 2$.

Let x_t denote the frequency of strategy R_1 at time step t and, similarly, $1 - x_t$ the frequency of strategy R_2. Assuming random pairwise interactions, the expected payoffs of R_1 and R_2 at time step t, denoted by $\pi_{1,t}$ and $\pi_{2,t}$, respectively, are given by

$$\begin{aligned} \pi_{1,t} &= x_t a_t + (1 - x_t) b_t \ , \\ \pi_{2,t} &= x_t c_t + (1 - x_t) d_t \ . \end{aligned} \tag{2.2.2}$$

The fitnesses of R_1-strategists and R_2-strategists at time step t are defined as $f_{1,t} = e^{\pi_{1,t}}$ and $f_{2,t} = e^{\pi_{2,t}}$, respectively (Vincent and Fisher, 1988; Blume, 1993; Tao *et al.*, 1997; Szabó and Hauert, 2002; Traulsen *et al.*, 2006; Claussen and Traulsen 2008; Traulsen *et al.*, 2008). Then, the average fitness of the population at time step t is $\bar{f}_t = x_t f_{1,t} + (1 - x_t) f_{2,t}$, and the frequency of R_1 at time step $t + 1$ can be expressed as

$$x_{t+1} \quad = \quad \frac{x_t e^{\pi_{1,t}}}{x_t e^{\pi_{1,t}} + (1 - x_t) e^{\pi_{2,t}}} \tag{2.2.3}$$

for $t \geq 0$ (Tao *et al.*, 1997; Traulsen *et al.*, 2008)

2.3 Random environmental noise and stochastic local stability

In the absence of random environmental noise, that is, in the situation where $\sigma_{ij}^2 = 0$ for all $i, j = 1, 2$ so that the payoff matrix in Eq. (2.2.1) is a constant matrix $\mathbf{A}(t) = \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix}$, Eq. (2.2.3) reduces to a deterministic recurrence equation. For this deterministic recurrence equation, it has been shown that: (*i*) only one equilibrium x^* with $0 < x^* < 1$ (called interior equilibrium) exists and is given by $x^* = (\bar{b} - \bar{d})/\gamma$ if $\bar{b} - \bar{d}$ and $\bar{c} - \bar{a}$ are both positive or both negative, where $\gamma = \bar{b} - \bar{d} + \bar{c} - \bar{a}$; (*ii*) x^* is globally asymptotically stable if $0 < \gamma < 2/x^*(1 - x^*)$; and (*iii*) as γ increases such that $\gamma > 2/x^*(1 - x^*)$, there are period-doubling bifurcation and chaos (Tao *et al.*, 1997). In the special case where $x^* = 1/2$, for instance, only one stable periodic two-cycle is possible for $\gamma > 8$.

If at least one σ_{ij}^2 for $i, j = 1$ or 2 is nonzero, which means that the random environmental noise is not degenerate, then Eq. (2.2.3) is a stochastic recurrence equation. In order to study the asymptotic (or long-run) behavior of the process $\{x_t\}$, suppose that $\tilde{x}$ is a constant (nonrandom) equilibrium of $\{x_t\}$, that is, an equilibrium of Eq. (2.2.3) that does not depend on the randomness of the payoff matrix. Obviously, both $\tilde{x} = 0$ and $\tilde{x} = 1$ are constant equilibria of Eq. (2.2.3) (called also the fixation states or the boundary equilibria of the system). Moreover, $\tilde{x}$ is called a constant interior equilibrium if it satisfies $0 < \tilde{x} < 1$ (Karlin and Liberman, 1974, 1975; Zheng *et al.*, 2017). Based on definitions introduced in Karlin and Liberman (1974, 1975), a constant equilibrium $\tilde{x}$ is said to be stochastically locally stable (SLS) if for any $\epsilon > 0$ there exists $\delta_0 > 0$ such that $\mathbb{P}\big(x_t \to \tilde{x}\big) \geq 1 - \epsilon$ as soon as $|x_0 - \tilde{x}| < \delta_0$, while a constant equilibrium $\tilde{x}$ is said to be stochastically locally unstable (SLU) if $\mathbb{P}\big(x_t \to \tilde{x}\big) = 0$ as soon as $|x_0 - \tilde{x}| > 0$.

Using the above definitions, we give below simplified mathematical arguments for the stochastic local stability of a constant equilibrium (the more rigorous mathematical proofs are similar to those in Zheng *et al.*, 2017.)

Let $u_t = x_t/(1-x_t)$. Then Eq. (2.2.3) can be equivalently expressed as

$$u_{t+1} \quad = \quad u_t e^{\pi_{1,t}-\pi_{2,t}} \ . \tag{2.3.1}$$

Consider first the stochastic local stability of the boundary equilibrium $\tilde{x} = 0$, which corresponds to $\tilde{u} = 0$. Notice that $u_t \to 0$ if and only if $x_t \to 0$. Iterating the above recurrence equation leads to

$$\begin{aligned} \frac{1}{n}\Big(\log u_n - \log u_0\Big) &= \frac{1}{n}\sum_{t=0}^{n-1}\big(\pi_{1,t} - \pi_{2,t}\big) \\ &= \frac{1}{n}\sum_{t=0}^{n-1}\left(\big(b_t - d_t\big) + \big(a_t - b_t - c_t + d_t\big)\frac{u_t}{1-u_t}\right) \ . \end{aligned} \tag{2.3.2}$$

Therefore, if $u_t \to 0$ and $\bar{b} - \bar{d} \neq 0$, then the strong law of large numbers guarantees that

$$0 \geq \lim_{n\to\infty}\frac{1}{n}\Big(\log u_n - \log u_0\Big) \quad \approx \quad \lim_{n\to\infty}\frac{1}{n}\sum_{t=0}^{n-1}\big(b_t - d_t\big) \ = \ \bar{b} - \bar{d} \ . \tag{2.3.3}$$

We conclude that the boundary equilibrium $\tilde{x} = 0$ is SLU if $\bar{b} - \bar{d} > 0$. On the other hand, using Egorov's theorem, it can be shown that $\tilde{x} = 0$ is SLS if $\bar{b} - \bar{d} < 0$ (Karlin and Liberman, 1975; Zheng *et al.*, 2017).

Suppose that there exists $\tilde{u} > 0$ such that

$$\tilde{u}\big(a_t - c_t\big) \quad = \quad d_t - b_t \tag{2.3.4}$$

for all possible $t \geq 0$. Then the random payoff matrix can be expressed as

$$\begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} \quad = \quad \begin{pmatrix} c_t + z_t & b_t \\ c_t & b_t + \tilde{u}z_t \end{pmatrix} \quad = \quad \begin{pmatrix} a_t & d_t - \tilde{u}z_t \\ a_t - z_t & d_t \end{pmatrix} , \tag{2.3.5}$$

where $z_t = a_t - c_t$. In this situation, Eq. (2.2.3) admits a constant interior equilibrium, given by $\tilde{x} = \tilde{u}/(1+\tilde{u})$ with $0 < \tilde{x} < 1$. Moreover, the previous analysis ascertains that the two fixation states $\tilde{x} = 0$ and $\tilde{x} = 1$ are both SLS if $\bar{z} = \bar{a} - \bar{c} > 0$ and both SLU if $\bar{z} = \bar{a} - \bar{c} < 0$. Now, in order to study the stochastic local stability of $\tilde{x} = \tilde{u}/(1+\tilde{u})$, let Eq. (2.3.1) be rewritten as

$$u_{t+1} \quad = \quad u_t e^{-\tilde{u}z_t + (1+\tilde{u})z_t\frac{u_t}{1+u_t}} \ . \tag{2.3.6}$$

A Taylor expansion around $\tilde{u}$ leads to the approximation

$$u_{t+1} - \tilde{u} \quad \approx \quad \left(1 + \frac{\tilde{u}}{1+\tilde{u}} z_t\right) \left(u_t - \tilde{u}\right) \ , \tag{2.3.7}$$

from which

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n}\Big(\log(u_n - \tilde{u})^2 - \log(u_0 - \tilde{u})^2\Big) \quad &\approx \quad \lim_{n\to\infty} \frac{1}{n} \log\left(1 + \frac{\tilde{u}}{1+\tilde{u}} z_t\right)^2 \\ &= \quad \left\langle \log\left(1 + \frac{\tilde{u}}{1+\tilde{u}} z_t\right)^2 \right\rangle \ . \end{aligned} \tag{2.3.8}$$

Then it can be shown that the constant interior equilibrium $\tilde{x} = \tilde{u}/(1+\tilde{u})$ is SLS if

$$\left\langle \log\left(1 + \tilde{x} z_t\right)^2 \right\rangle \quad < \quad 0 \ , \tag{2.3.9}$$

and SLU if the inequality is reversed. Developing $\log\left(1 + \tilde{x} z_t\right)^2$ around $\bar{z} = \bar{a} - \bar{c}$ yields

$$\left\langle \log\left(1 + \tilde{x} z_t\right)^2 \right\rangle \quad \approx \quad \log\left(1 + \tilde{x}\bar{z}\right)^2 - \left(\frac{\tilde{x}}{1+\tilde{x}\bar{z}}\right)^2 \sigma_z^2 \ , \tag{2.3.10}$$

where $\sigma_z^2 = \sigma_a^2 + \sigma_c^2 - 2\sigma_{a,c}$. Therefore, as long as the random environmental noise does not vary too much, the constant interior equilibrium $\tilde{x} = \tilde{u}/(1+\tilde{u})$ is SLS if

$$\sigma_z^2 \quad > \quad \left(\frac{1+\tilde{x}\bar{z}}{\tilde{x}}\right)^2 \log\left(1 + \tilde{x}\bar{z}\right)^2 \ , \tag{2.3.11}$$

and SLU if the inequality is reversed. This result shows not only that the two boundary equilibria ($\tilde{x} = 0$ and $\tilde{x} = 1$) and the constant interior equilibrium ($\tilde{x} = \tilde{u}/(1+\tilde{u})$) can be simultaneously SLS, but also that an increase in the variance of the environmental noise (σ_z^2) will promote the stochastic local stability of the constant interior equilibrium.

In order to test the above theoretical predictions, three numerical examples are investigated below using computer simulations in the case of a random payoff matrix in the form of Eq. (2.3.6) with a constant interior equilibrium $\tilde{x} = \tilde{u}/(1+\tilde{u})$.

2.3.1 Example 1

If we take $\tilde{u} = 1$ in Eq. (2.3.6), then $\tilde{x} = 1/2$ is a constant interior equilibrium. From our theoretical results, both $\tilde{x} = 0$ and $\tilde{x} = 1$ are SLS if $\bar{z} > 0$, while $\tilde{x} = 1/2$ is SLS if

$\sigma_z^2 > \left(2+\bar{z}\right)^2 \log\left(1+\bar{z}/2\right)^2$. The simulation results based on Eq. (2.2.3) are plotted in Figure 2.1, in which we take $\bar{z} = 0.1$ and $\sigma_z^2 = 4$. These simulations strongly support the theoretical predictions, that is, both boundaries and the constant interior equilibrium can be simultaneously SLS. Notice that in the absence of random environmental noise (σ_z^2), both boundaries and the constant interior equilibrium cannot be simultaneously locally stable.

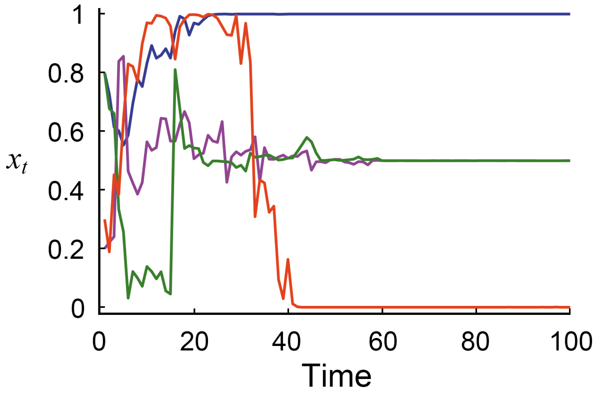


Figure 2.1: *Both boundaries and the constant interior equilibrium can be stochastically locally stable at the same time.* For the random payoff matrix $\begin{pmatrix} 1+z_t & 1 \\ 1 & 1+z_t \end{pmatrix}$ with $\bar{z} = 0.1$ and $\sigma_z^2 = 4$, both boundaries and $\tilde{x} = 1/2$ are SLS. The simulation results, four trajectories of x_t, are illustrated starting with $x_0 = 0.2$ and $x_0 = 0.8$: two converge to $\tilde{x} = 1/2$ (green and pink curves), one to $\tilde{x} = 0$ (red curve), and one to $\tilde{x} = 1$ (blue curve).

2.3.2 Example 2

Similarly to *Example 1*, we take $\tilde{u} = 1$ in Eq. (2.3.6) so that $\tilde{x} = 1/2$ is a constant interior equilibrium. If $\bar{z} < 0$, then both $\tilde{x} = 0$ and $\tilde{x} = 1$ are SLU. On the other hand, if $\sigma_z^2 = 0$, then $\tilde{x} = 1/2$ is globally asymptotically stable if $|\bar{z}| < 4$ (with $\bar{z} < 0$), while only one stable periodic two-cycle can exist when $|\bar{z}| > 4$ (Tao *et al.*, 1997) (see Figure 2.2a). For

$\bar{z} = -6$ and $\sigma_z^2 > 0$, the simulation results show the following: (i) when σ_z^2 is small, the probability distribution of x_t over time is bimodal about $\tilde{x} = 1/2$ (see Figure 2.2b); and (ii) as σ_z^2 increases, a new peak of probability distribution appears at $\tilde{x} = 1/2$, and the amount of probability near the constant interior equilibrium $\tilde{x} = 1/2$ rapidly increases (see Figure 2.2c-d). Obviously, the increase of σ_z^2 promotes the stochastic local stability of $\tilde{x} = 1/2$. However, we can see also that, although the increase of σ_z^2 leads to a new peak of the probability distribution at $\tilde{x} = 1/2$, the period doubling characteristic of the system is not completely destroyed if σ_z^2 is not too large.

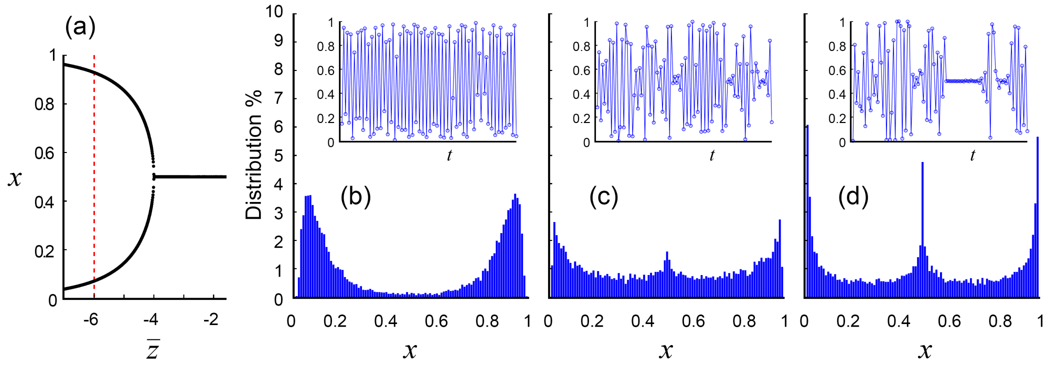


Figure 2.2: *An increase of σ_z^2 promotes stochastic local stability of $\tilde{x} = 1/2$.* (**a**) For the random payoff matrix $\begin{pmatrix} 1+z_t & 1 \\ 1 & 1+z_t \end{pmatrix}$ with $\bar{z} = -6$, if $\sigma_z^2 = 0$, both boundaries and the constant interior equilibrium $\tilde{x} = 1/2$ are unstable and there is a stable periodic two-cycle (red dash line). (**b**) For $\sigma_z^2 = 1$, the time evolution of x_t corresponds to a bimodal probability distribution about $\tilde{x} = 1/2$, and the system still has the obvious two-cycle characteristics. (**c-d**) For $\sigma_z^2 = 4$ in panel (**c**) and $\sigma_z^2 = 16$ in panel (**d**), we can see that increasing σ_z^2 results in a new peak of probability distribution at $\tilde{x} = 1/2$ and an increase in the amount of probability near $\tilde{x} = 1/2$.

2.3.3 Example 3

In this example, we take $\tilde{u} = 1/2$ in Eq. (2.3.6) so that $\tilde{x} = 1/3$ is a constant interior equilibrium. If $\bar{z} < 0$, then both $\tilde{x} = 0$ and $\tilde{x} = 1$ are SLU. On the other hand, if $\sigma_z^2 = 0$, then $\tilde{x} = 1/3$ is globally asymptotically stable if $|\bar{z}| < 6$ (with $\bar{z} < 0$), while an increase of $|\bar{z}|$ (with $|\bar{z}| > 6$) leads to period-doubling bifurcation and chaos (Tao *et al.*, 1997) (see Figure 2.3a). Here, we take $\bar{z} = -9$ so that the system exhibits a stable periodic four-cycle if $\sigma_z^2 = 0$ (see also Figure 2.3a). Similarly to the results in *Example 2*, we notice that: (i) when σ_z^2 is small (but $\sigma_z^2 \neq 0$), the probability distribution of x_t over time shows four peaks (this phenomenon exactly matches the nonlinear dynamical characteristics of the system) (see Figure 2.3b); and (ii) with the increase of σ_z^2, a new peak of the probability distribution appears at $\tilde{x} = 1/3$ and the amount of probability near $\tilde{x} = 1/3$ is also positively related to the size of σ_z^2 (see Figure 2.3c-f).

2.4 Conclusion and discussion

Our theoretical results and simulations on evolutionary games with a random payoff matrix clearly show that stochastic fluctuations in the payoffs as a result of random noise in the environment make the dynamical system much more complex, namely that an increase in the level of environmental noise could promote stochastic local stability of a constant interior equilibrium. Although it may look at a first glance that the matrix games considered in this letter have totally uncertain outcomes, this is not the case. The payoffs in pairwise interactions are not assumed to be independent identically distributed random variables. In particular, they are not assumed to have the same expected value. The structure of the game is determined by the expected payoffs which in turn determine the dynamical properties of fixation states as well as interior equilibria if they exist in the absence of variability. The local stability properties and conditions have to be extended when random perturbations on the payoffs are introduced. These extensions for fixation states can be used to define stochastic evolutionary concepts such as stochastic evolutionary stability

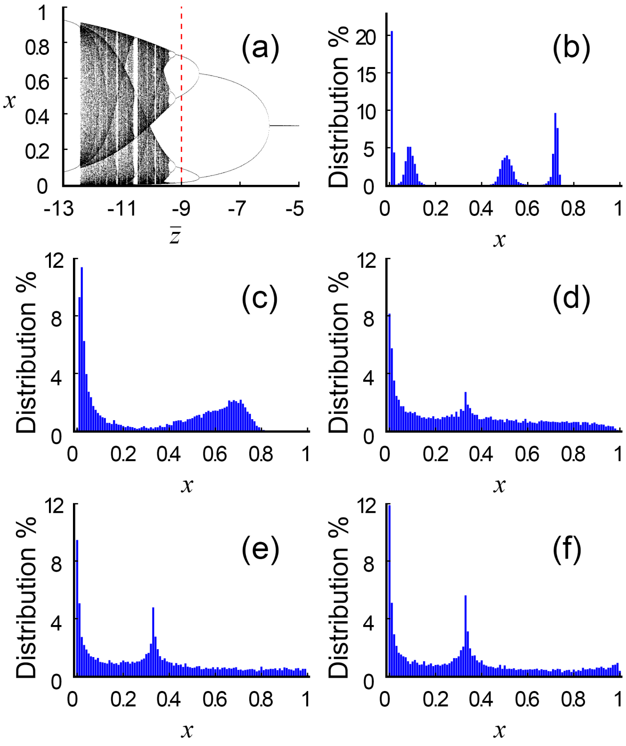


Figure 2.3: *An increase of σ_z^2 promotes the stochastic local stability of $\tilde{x} = 1/3$.* (**a**) For the random payoff matrix $\begin{pmatrix} 1+z_t & 1 \\ 1 & 1+z_t/2 \end{pmatrix}$ with $\bar{z} = -9$, if $\sigma_z^2 = 0$, both boundaries and the constant interior equilibrium $\tilde{x} = 1/3$ are unstable and there is a stable periodic four-cycle (red dash line). (**b**) For $\sigma_z^2 = 0.01$, the probability distribution of x_t exhibits four peaks, caracteristic of a stable periodic four-cycle, when σ_z^2 is small (but $\sigma_z^2 \neq 0$). (**c**-**f**) For $\sigma_z^2 = 0.25$ in panel (**c**), $\sigma_z^2 = 9$ in panel (**d**), $\sigma_z^2 = 16$ in panel (**e**) and $\sigma_z^2 = 25$ in panel (**f**), we can see that increasing σ_z^2 not only leads to the appearance of a new peak of probability distribution of x_t at $\tilde{x} = 1/3$, but also to an increase in the amount of probability near $\tilde{x} = 1/3$.

and stochastic convergence stability (Zheng *et al.*, 2017).

In this letter, we have focused on the existence of a stochastically locally stable interior equilibrium in a discrete-time two-phenotype model with an exponential function of the payoff as fitness to best approximate a continuous-time model. Notice that our analysis is fully valid if we multiply all payoffs by a common positive factor $\beta = 1/c$ that would represent the strength of selection (Traulsen *et al.*, 2008). However, the focus in this letter is not on the effect of such a parameter and, in order to keep the number of parameters as low as possible without loss of generality, the strength of selection will be incorporated into the payoffs.

The three examples studied in the paper allow us to address the global dynamics of the system and make evolutionary predictions in the most interesting cases, namely, when the fixation states are both stochastically locally stable (SLS) or both stochastically locally unstable (SLU). As shown, an increase in the variance of environmental noise (σ_z^2) favors the stochastic local stability of an interior equilibrium even in the former case, which is rather surprising. In our simulations, a gradual increase in the environmental noise intensity leads to a gradual increase in the probability distribution of the population state over a long period of time near the constant interior equilibrium at which all individuals have the same average fitness. Moreover, this is in agreement with the mathematical condition for a constant interior equilibrium (Eq. (2.3.11)) to be SLS, so that the population state tends to wander around it. Notice that the stochastic local stability of the constant interior equilibrium depends not only on the averages of the payoffs but also on their variances and covariances contrary to the boundary equilibria. This is a characteristic of the stochastic model compared to the deterministic model that can make possible the coexistence of a SLS interior equilibrium with two SLS boundary equilibria as environmental noise intensity increases.

Our conclusion may seem counterintuitive and have important biological implications. There is a connection, however, between our results on SLS equilibria in an evolutionary game model and ideas of noise-induced transitions in stochastic differential equations models in physics, chemistry and biology (Horsthemke and Lefever, 2006). Random fluctuations in the environment correspond to what is called multiplicative noise, and it is known that the number of peaks in the stationary probability density of a diffusion process can change as a function of noise intensity. In a population genetics context, a diffusion process with a stationary distribution can be obtained as a continuous-time limit of a discrete-time mutation-selection model. In our model, there is no mutation, selection is frequency-dependent and strong, and time is kept discrete. Nevertheless, there is an analogy between a SLS equilibrium in our model and a peak in the stationary probability density of a diffusion approximation, if it can be found, that would be obtained under weak mutation and weak frequency-dependent selection. In this case, however, a boundary SLS equilibrium would correspond to a peak not exactly on the boundary but nearby the boundary.

May (1976) found that a simple deterministic logistic difference equation can lead to periodic limit cycles and chaos. This discovery led people to believe that nonlinear biological systems could result in the emergence of complex dynamics, and that such dynamics (especially chaos) should be easily observable in natural populations. Nevertheless, the majority of attempts to find chaos in nature have either drawn a blank or remained controversial. Since then, several studies have aimed to explaining why natural populations do not exhibit chaos (Sherratt and Wilkinson, 2009). Our results on the effects of random noise on evolutionary game dynamics can provide some clues for addressing this question, mainly, that stochastic fluctuations in the environment may play a role in impeding the emergence of complex dynamical behaviors in natural populations. The explanation might be that a random environment favors the evolution of more robust equilibrium population strategies but this remains to be confirmed by further studies.

Acknowledgements

In this study, X.-D.Z. and Y.T. were supported by the National Natural Science Foundation of China (Grants No.31770426, No.11471311 and No.11401562) and the National Basic Research Program of China (Grant No.2013CB945000); S.L. was supported by Chinese Academy of Sciences President's International Fellowship Initiative (Grant No. 2016VBA039); C.L. and S.L. were supported in part by Natural Sciences and Engineering Research Council of Canada, Grant No.8833.

X.-D.Z. and C.L. contributed equally to this work.

Bibliography

[1] J. Maynard Smith, G.R. Price, The logic of animal conflicts. *Nature* **246**, 15-18 (1973).

[2] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, UK, 1982).

[3] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[4] S. Lessard, Evolutionary dynamics in frequency-dependent two-phenotype models. *Theoretical Population Biology* **25**, 210-234 (1984).

[5] J. Hofbauer, K. Sigmund, *The Theory of Evolution and Dynamical Systems* (Cambridge University Press, Cambridge, UK, 1988).

[6] M.A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).

[7] M. Broom, J. Rychtář, *Game-Theoretical Models in Biology* (CRC Press, London, UK, 2013)

[8] P.D. Taylor, L. Jonker, Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* **40**, 145-156 (1978).

[9] E.C. Zeeman, in *Global Theory of Dynamical Systems* (Springer Verlag, Berlin, Heidelberg, 471, 1980).

[10] J.W. Weibull, *Evolutionary Game Theory* (MIT Press, Cambridge, MA, 1995).

[11] G. Szabó, C. Hauert, Phase transitions and volunteering in spatial public goods games. *Physical Review Letters* **89**, 118101 (2002).

[12] A. Traulsen, M.A. Nowak, and J.M. Pacheco, Stochastic dynamics of invasion and fixation. *Physical Review E* **74**, 011909 (2006).

[13] J.C. Claussen and A. Traulsen, Cyclic dominance and biodiversity in well-mixed populations. *Physical Review Letters* **100**, 058104 (2008).

[14] A. Traulsen, N. Shoresh and M.A. Nowak, Analytical Results for Individual and Group Selection of Any Intensity. *Bulletin of Mathematical Biology* **70**, 1410-1424 (2008).

[15] Y. Tao, Q.-S. Yang, Z.-G. Jiang and Z.-W. Wang, Evolutionarily stable strategy, stable state, periodic cycle and chaos in a simple discrete time two-phenotype model. *Journal of Theoretical Biology* **188**, 21-27 (1997).

[16] T.L. Vincent and M.E. Fisher, Evolutionarily stable strategies in differential and difference equation models. *Evolutionary Ecology* **2**, 321, (1988).

[17] L.E. Blume, The statistical mechanics of strategic interaction. *Games and Economic Behavior* **5**, 387-424 (1993).

[18] R. M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, NJ, 1973).

[19] P. Turchin, L. Oksanen, P. Ekerholm, T. Oksanen, H. Henttonen, Are lemmings prey or predators? *Nature* **405**, 562 (2000); O.N. Bjornstad, B.T. Grenfell, Noisy clockwork: Time series analysis of population fluctuations in animals *Science* **293**, 638-642 (2001); X.-D. Zheng, L.-L. Deng, W.-Y. Qiang, R. Cressman, Y. Tao, Limiting similarity of competitive species and demographic stochasticity. *Physical Review E* **95**, 042404 (2017).

[20] R. Lande, S. Engen, B.-E. Sæther, *Stochastic Population Dynamics in Ecology and Conservation* (Oxford University Press, 2003).

[21] X.-D. Zheng, C. Li, S. Lessard, Y. Tao, Evolutionary stability concepts in a stochastic environment. *Physical Review E* **96**, 032414 (2017).

[22] S. Karlin, U. Liberman, Random temporal variation in selection intensities: case of large population size. *Theoretical Population Biology* **6**, 355-382 (1974).

[23] S. Karlin, U. Liberman, Random temporal variation in selection intensities: One-locus two-allele model. *Journal of Mathematical Biology* **2**, 1-17 (1975).

[24] W. Horsthemke, R. Lefever, *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry , and Biology* (Springer-Verlag, 2006).

[25] R.M. May, Simple mathematical models with very complicated dynamics. *Nature* **261**, 459-467 (1976).

[26] T.N. Sherratt, D.M. Wilkinson, *Big Questions in Ecology and Evolution* (Oxford University Press, 2009).

Chapter 3

Article 3

Weak selection can filter environmental noise in the evolution of animal behavior

Cong Li[1,2,*], Xiu-Deng Zheng[1,*], Tian-Jiao Feng[3], Sabin Lessard[2] and Yi Tao[1,4]

[1]Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

[2] Department of Mathematics and Statistics, University of Montreal, Montreal QC H3C 3J7, Canada

[3] School of Systems Science, Beijing Normal University, Beijing 100875, China

[4]University of Chinese Academy of Sciences, Beijing 100049, China

[*]These authors contributed equally to this work.

Information

This paper appeared in *Physical Review E* **100**, 052411. It was received on 2 January 2019; revised on 4 September 2019; published on 26 November 2019.

My contribution

I did most of the theoretical analysis and revised the paper based on a draft by Yi Tao. The motivation of this article comes from the discussion of the examples in Article 4. I worked on the example for Figure 3.1 and ran the numerical simulations.

Abstract

Weak selection is an important assumption in theoretical evolutionary biology, but its biological significance remains unclear. In this study, we investigate the effect of weak selection on stochastic evolutionary stability in a two-phenotype evolutionary game dynamics with a random payoff matrix assuming an infinite, well-mixed population undergoing discrete, nonoverlapping generations. We show that, under weak selection, both stochastic local stability and stochastic evolutionary stability in this system depend on the means of the random payoffs but not on their variances. Moreover, although stochastic local stability or instability of an equilibrium may not depend on environmental noise if selection is weak enough, the growth rate near an equilibrium not only depends on environmental noise, but it can even be enhanced by environmental noise if selection is weak. This is the case, for instance, when the variances of the random payoffs are equal as well as the covariances. These results suggest that natural selection could be able to filter (or resist) the effect of environmental noise on the evolution of animal behavior if selection is weak.

3.1 Introduction

Weak selection is an important assumption in theoretical evolutionary biology. It is the assumption that there is little difference between the individuals in reproductive success, or fitness, so that the effects of natural selection are small. Weak selection has a long-standing history in population genetics (Kimura, 1968; Ohta, 2002). In infinitely large populations in a constant environment, however, increasing the intensity of selection often results in a mere re-scaling of time which does not actually affect the final outcome of the deterministic dynamics (Hofbauer and Sigmund, 1998; Traulsen *et al.*, 2005). On the opposite, in finite populations, changing the intensity of selection may have an important effect on the stochastic dynamics (Kimura, 1968; Nowak *et al.*, 2004). In some situations, results under weak selection have been shown to stay valid as the intensity of selection increases (Ohtsuki *et al.*, 2006). In general, however, the evolutionary significance of weak selection in finite populations remains unclear.

The assumption of weak selection has already been considered in evolutionary game theory to analyze the stochastic dynamics in finite populations (Nowak, 2006). Here, weak selection means that the expected payoff of an individual has only a very small effect on its fitness so that the evolutionary dynamics is mainly driven by random fluctuations (Nowak *et al.*, 2004; Taylor *et al.*, 2004). Under the assumption of weak selection, Nowak *et al.* (2004) deduced the “one-third law” for the fixation probability in a two-phenotype game-theoretic model and used it to provide an explanation for the evolution of cooperation (see also Lessard, 2005; Traulsen *et al.*, 2006a, 2006b; Lessard and Ladret, 2007; Traulsen and Hauert, 2009; Zheng *et al.*, 2011). In order to show the robustness of outcomes in finite populations under weak selection, Wu *et al.* (2010, 2013) investigated some properties of weak selection in the Fermi and Moran processes, where the environment is assumed to be fixed so that the payoff matrix remains constant.

Environmental conditions in the real world are changing and uncertain, and stochastic fluctuations in the surroundings of a population may cause changes in the occurrence of interactions between individuals and, more importantly, changes in the payoffs received by the interacting individuals (Zheng *et al.*, 2017, 2018). As pointed out by May (1973), the birth rates, carrying capacities, competition coefficients, and other ecological parameters which characterize natural biological systems all, to a greater or lesser degree, exhibit random fluctuations. Therefore, a very challenging question is whether natural selection is able to filter (or resist) the effect of environmental noise on the evolution of animal behavior.

Recently, in order to develop the concept of evolutionary stability in a randomly fluctuating environment, Zheng *et al.* (2017, 2018) investigated conditions for stochastic local stability of the fixation states and constant interior equilibria in a two-phenotype model with random payoffs, and developed the concepts of stochastic evolutionary stability and stochastic convergence stability. The results obtained show that stochastic local stability depends not only on the averages of the random payoffs but also on the variances of these random payoffs. Note that Stollmeier and Nagler (2018) considered also an evolutionary game dynamics with two phenotypes and time-dependent payoffs in an infinite population undergoing discrete, nonoverlapping generations, but they focused on the unfair coexistence of strategies.

Extending the analysis of stochastic local stability and stochastic evolutionary stability, we are interested in this paper in what determines the characteristics of the evolutionary game dynamics in the presence of environmental noise if selection is weak. Our main goal is to reveal the fundamental importance of weak selection in the evolution of animal behavior, or the evolutionary biological significance of weak selection, in a stochastic environment.

It may be useful to recall that stochastic fluctuations in evolutionary game dynamics may be due to either intrinsic noise (i.e., demographic stochasticity) or extrinsic noise (i.e., environmental stochasticity), or a combination of both. Demographic stochasticity mainly involves the occurrence of interactions between individuals, random events of birth and death of individuals, etc. Demographic stochasticity in evolutionary game dynamics due to a finite population size has received a lot of attention as already mentioned (Nowak *et al.*, 2004; Taylor *et al.*, 2004; Lessard, 2005; Ohtsuki *et al.*, 2006; Nowak, 2006; Traulsen *et al.*, 2006a, 2006b; Lessard and Ladret, 2007; Traulsen and Hauert, 2009; Zheng *et al.*, 2011). On the other hand, stochastic fluctuations in the population state due to a finite population size can be much smaller than those caused by changes in the environment, and then ignored, if the population size is large enough. This assumption is current in evolutionary game theory (Maynard Smith, 1982; Lessard, 1984; Hofbauer and Sigmund, 1998), and deserves as much attention as the assumption of a population size whose inverse is of order larger than, or equal to, the order of random differences in payoffs. Weak selection, however, another current assumption in evolutionary game theory, may come into play in the short-term as well as long-term effects of random fluctuations in the environment. This is the question addressed in the present paper, which was not addressed in previous studies.

3.2 Basic model and definitions

Consider an evolutionary game in an infinite population with discrete, nonoverlapping, generations. There are two phenotypes or pure strategies, S_1 and S_2, and the payoffs in pairwise interactions at time step $t \geq 0$ are given by the game matrix

$$\mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix} = \begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} , \tag{3.2.1}$$

where $a_{ij}(t)$ is the payoff to strategy S_i against strategy S_j for $i, j = 1, 2$. These payoffs are assumed to be positive random variables that are uniformly bounded below and above by some positive constants. Therefore, there exist real numbers $A, B > 0$ such that $A \leq a_{ij}(t) \leq B$ for $i, j = 1, 2$ and all $t \geq 0$ (Zheng *et al.*, 2017). Moreover, the probability distributions of $a_{ij}(t)$ for $i, j = 1, 2$ do not depend on $t \geq 0$. The means, variances and covariances of these random payoffs are given by $\left\langle a_{ij}(t) \right\rangle = \bar{a}_{ij}$, $\left\langle \left(a_{ij}(t) - \bar{a}_{ij}\right)^2 \right\rangle = \sigma_{ij}^2$, and $\left\langle \left(a_{ij}(t) - \bar{a}_{ij}\right)\left(a_{kl}(t) - \bar{a}_{kl}\right) \right\rangle = \sigma_{ij,kl}$, respectively, for $i, j, k, l = 1, 2$ with $(i, j) \neq (k, l)$.

As for $s \neq t$, the payoffs $a_{ij}(s)$ and $a_{kl}(t)$ are assumed to be independent of each other so that $\Big\langle \big(a_{ij}(s) - \bar{a}_{ij}\big)\big(a_{kl}(t) - \bar{a}_{kl}\big) \Big\rangle = 0$ for $i, j, k, l = 1, 2$. In general, we also assume that the variances of the random payoffs are small (Zheng *et al.*, 2017, 2018)

Let x_t be the frequency of strategy S_1 at time step $t \geq 0$ and, similarly, $1 - x_t$ the frequency of strategy S_2. Then the expected payoffs of strategies S_1 and S_2 at time step $t \geq 0$ are given by $\pi_{1,t} = x_t a_t + (1 - x_t) b_t$ and $\pi_{2,t} = x_t c_t + (1 - x_t) d_t$, respectively. Furthermore, in order to show the effect of selection intensity on the evolutionary dynamics of strategies S_1 and S_2, and without loss of generality, the fitnesses of S_1 and S_2 at time step $t \geq 0$ are simply defined as $(1 - w) + w\pi_{1,t}$ and $(1 - w) + w\pi_{2,t}$, respectively, where w with $0 \leq w \leq 1$ represents the selection intensity (Nowak *et al.*, 2004; Nowak, 2006). So, the number of replicates of a strategy from one step to the next is proportional to its fitness, and the frequency of strategy S_1 at time step $t + 1$ is given by the recurrence equation

$$x_{t+1} = \frac{x_t\big((1 - w) + w\pi_{1,t}\big)}{x_t\big((1 - w) + w\pi_{1,t}\big) + (1 - x_t)\big((1 - w) + w\pi_{2,t}\big)} \tag{3.2.2}$$

for $t \geq 0$. This model can be viewed as a Wright-Fisher model in the limit of a large population size (see, e.g., Hofbauer and Sigmund, 1998), but with fitness differences of order larger than the inverse of the population size and subject to stochastic fluctuations. Defining $u_t = x_t \big/ (1 - x_t)$, the recurrence equation takes the simple form

$$u_{t+1} = u_t \left[\frac{u_t\big((1 - w) + wa_t\big) + \big((1 - w) + wb_t\big)}{u_t\big((1 - w) + wc_t\big) + \big((1 - w) + wd_t\big)} \right] . \tag{3.2.3}$$

Let $\hat{x}$ represent a constant (non-random) equilibrium of Eq. (3.2.2) that does not depend on the randomness of the payoff matrix $\mathbf{A}(t)$. This is clearly the case for both $\hat{x} = 0$ and $\hat{x} = 1$, called the *fixation states* or the *boundary equilibria*. This may also be the case for a constant equilibrium $\hat{x}$ with $0 < \hat{x} < 1$, called a constant interior equilibrium. A constant equilibrium $\hat{x}$ is said to be *stochastically locally stable* (SLS) if for every $\epsilon > 0$ there exists $\delta_0 > 0$ such that $\mathbb{P}(x_t \to \hat{x}) \geq 1 - \epsilon$ as soon as $|x_0 - \hat{x}| < \delta_0$ (Karlin and Liberman, 1974, 1975; Zheng *et al.*, 2017). This means that x_t tends to $\hat{x}$ as $t \to \infty$ with probability arbitrarily close to 1 (but different from 1) if the initial state x_0 is sufficiently near $\hat{x}$. On the other hand, a constant equilibrium $\hat{x}$ can be said to be *stochastically locally unstable* (SLU) if $\mathbb{P}(x_t \to \hat{x}) = 0$ as soon as $|x_0 - \hat{x}| > 0$ (Karlin and Liberman, 1974, 1975; Zheng *et al.*, 2017). If this is the case, then $\hat{x}$ cannot be reached with probability 1 from any initial state different from $\hat{x}$. Based on these definitions, we will present some simplified mathematical arguments for the stochastic local stability of a constant equilibrium (the more rigorous mathematical proofs are similar to those in Zheng *et al.*, 2017

3.3 Effect of weak selection on stochastic local stability of an equilibrium

Consider first the stochastic local stability of the fixation state $\hat{x} = 0$ in Eq. (3.2.2), which corresponds to the equilibrium $\hat{u} = \hat{x}/(1 - \hat{x}) = 0$ in Eq. (3.2.3). Note that Eq. (3.2.3) can be rewritten in the form

$$\frac{u_{t+1}}{u_t} = \left[\frac{(1-w) + wb_t}{(1-w) + wd_t}\right] R_t \ , \tag{3.3.1}$$

where

$$R_t = 1 + \frac{u_t\big[\big((1-w)+wa_t\big)\big((1-w)+wd_t\big) - \big((1-w)+wb_t\big)\big((1-w)+wc_t\big)\big]}{u_t\big((1-w)+wb_t\big)\big((1-w)+wc_t\big) - \big((1-w)+wb_t\big)\big((1-w)+wd_t\big)} \ . \tag{3.3.2}$$

Then, iterating this recurrence equation leads to

$$\frac{1}{n}\Big[\log u_n - \log u_0\Big] = \frac{1}{n}\sum_{t=0}^{n-1} \log\left[\frac{(1-w) + wb_t}{(1-w) + wd_t}\right] + \frac{1}{n}\sum_{t=0}^{n-1} \log R_t \tag{3.3.3}$$

for $n \geq 1$. Therefore, if $u_t \to 0$ (which compels $\log R_t \to 0$), then the strong law of large numbers guarantees that

$$\lim_{n\to\infty} \frac{1}{n}\Big[\log u_n - \log u_0\Big] \approx \left\langle \log\left[\frac{(1-w) + wb_t}{(1-w) + wd_t}\right] \right\rangle \ . \tag{3.3.4}$$

Using Egorov's theorem, it can be shown that the fixation state $\hat{x} = 0$ is SLS if

$$\left\langle \log\left[\frac{(1-w)+wb_t}{(1-w)+wd_t}\right]\right\rangle = \left\langle \log\left[(1-w)+wb_t\right]\right\rangle - \left\langle \log\left[(1-w)+wd_t\right]\right\rangle < 0\ , \quad (3.3.5)$$

and $\hat{x} = 0$ is SLU if the inequality is reversed (Zheng *et al.*, 2017). The mean geometric growth rate on the left-hand side in Eq. (3.3.5) represents the rate of convergence to 0 if 0 is SLS and the rate of divergence from 0 if 0 is SLU.

In the case where the payoffs have small enough variances, we have the approximation

$$\left\langle \log\left[(1-w)+wa_{ij}(t)\right]\right\rangle \approx \log\left[(1-w)+w\bar{a}_{ij}\right] - \frac{w^2\sigma_{ij}^2}{2\big((1-w)+w\bar{a}_{ij}\big)^2} \quad (3.3.6)$$

for $i, j = 1, 2$ (Zheng *et al.*, 2017). Then, the inequality in Eq. (3.3.5) can be rewritten as

$$\log\left[\frac{(1-w)+w\bar{b}}{(1-w)+w\bar{d}}\right] + \frac{w^2\sigma_d^2}{2\big((1-w)+w\bar{d}\big)^2} - \frac{w^2\sigma_b^2}{2\big((1-w)+w\bar{b}\big)^2} < 0\ . \quad (3.3.7)$$

Furthermore, when w is small enough, we have the approximation

$$\log\left[\frac{(1-w)+w\bar{b}}{(1-w)+w\bar{d}}\right] \approx w\big(\bar{b}-\bar{d}\big)\ . \quad (3.3.8)$$

Therefore, if selection is weak enough, then the fixation state $\hat{x} = 0$ is SLS if $\bar{b} - \bar{d} < 0$ and SLU if $\bar{b} - \bar{d} > 0$. This implies that the stochastic local stability of $\hat{x} = 0$ depends on the means of the random payoffs b_t and d_t, but does not depend on their variances. An example of stochastic local stability of fixation state $\hat{x} = 0$ under weak selection is shown in **Fig. 3.1**. By symmetry, under weak enough selection, the fixation state $\hat{x} = 1$ is SLS if $\bar{c} - \bar{a} < 0$ and SLU if $\bar{c} - \bar{a} > 0$. On the other hand, in the degenerate case where $b_t = d_t$ (or $a_t = c_t$) for all $t \geq 0$, and under weak enough selection, the fixation state $\hat{x} = 0$ (or $\hat{x} = 1$) is SLS if $\bar{a} - \bar{c} < 0$ (or $\bar{d} - \bar{b} < 0$) and SLU if $\bar{a} - \bar{c} > 0$ (or $\bar{d} - \bar{b} > 0$). (The mathematical proofs are given in **Appendix A**.)

Moreover, as a special case, if $\hat{u}(a_t - c_t) = d_t - b_t$ for all $t \geq 0$ where $\hat{u}$ is a positive constant, then the random payoff matrix $\mathbf{A}(t)$ in Eq. (3.2.1) can be re-written as

$$\begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} = \begin{pmatrix} c_t + z_t & b_t \\ c_t & b_t + \hat{u}z_t \end{pmatrix} = \begin{pmatrix} a_t & d_t - \hat{u}z_t \\ a_t - z_t & d_t \end{pmatrix}\ , \quad (3.3.9)$$

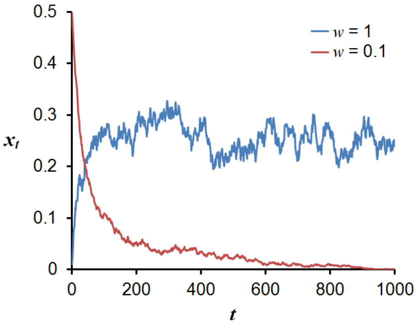


Figure 3.1: *Effect of selection intensity on the stochastic local stability of fixation state* $\hat{x} = 0$*. Consider a random payoff matrix* $\mathbf{A}(t) = \begin{pmatrix} 7 & 9+\eta_t \\ 8 & 10+\xi_t \end{pmatrix}$*, where* η_t *and* ξ_t *are* uniform random variables with $\bar{\eta} = \bar{\xi} = 0$, $\sigma_\eta^2 = 5.3$ and $\sigma_\xi^2 = 30$ for all $t \geq 0$. Simulation results illustrate the stochastic local stability or instability of $\hat{x} = 0$ for two different intensities of selection. When $w = 1$, then $\hat{x} = 0$ is SLU and the population state is driven away from 0 even from an initial state close to 0 such as $x_0 = 0.01$. When $w = 0.1$, then $\hat{x} = 0$ is SLS and the population state tends to 0. Each curve represents an average of 100 simulated trajectories starting from the same initial state. Notice that each trajectory in the case $w = 1$ fluctuates between 0 and 1 without any convergence.

where $z_t = a_t - c_t$ for all $t \geq 0$. For this random payoff matrix, $\hat{x} = \hat{u}/(1+\hat{u})$ is a constant interior equilibrium of Eq. (3.2.2). Similarly to the the stochastic local stability analysis of the fixation state $\hat{x} = 0$, it can be shown that under weak selection, the constant interior equilibrium $\hat{x} = \hat{u}/(1+\hat{u})$ is SLS if $\bar{c} - \bar{a} > 0$ and SLU if $\bar{c} - \bar{a} < 0$ (the mathematical proofs are given in **Appendix B**). This result shows that, if a constant interior equilibrium exists, then its stochastic local stability under weak selection depends on the means of the random payoffs but not on their variances. However, we have to point out that even if selection is weak, whether a constant interior equilibrium exists or not cannot be in general determined only by the means of the random payoffs.

3.4 Effect of environmental noise on the growth rate near an equilibrium under weak selection

A further challenging question concerns the rate of convergence (or divergence) near an equilibrium in Eq. (3.2.2) with the random payoff matrix $\mathbf{A}(t) = \big(a_{ij}(t)\big)_{2\times 2}$ in Eq. (3.2.1) at time step $t \geq 0$, compared to the deterministic dynamics with the constant mean payoff matrix $\bar{\mathbf{A}} = \big(\bar{a}_{ij}\big)_{2\times 2}$: does this rate increase or decrease as the variance in the payoffs increases?

Consider first the situation where the fixation state $\hat{x} = 0$ is SLS in the stochastic dynamics under weak selection. Owing to Eq. (3.3.7) and Eq. (3.3.8), the rate of convergence to 0 is approximated as

$$\left\langle \log\left[\frac{(1-w)+wb_t}{(1-w)+wd_t}\right]\right\rangle \approx \begin{cases} w(\bar{b}-\bar{d}) + \frac{w^2\left(\sigma_d^2-\sigma_b^2\right)}{2} & \text{if } \ \sigma_b^2 \neq \sigma_d^2 \ , \\ w(\bar{b}-\bar{d}) + w^3\sigma^2\big(\bar{b}-\bar{d}\big) & \text{if } \ \sigma_b^2 = \sigma_d^2 = \sigma^2 \ , \end{cases} \tag{3.4.1}$$

where $w(\bar{b}-\bar{d}) < 0$ approximates the rate of convergence in the deterministic mean-field dynamics with payoff matrix $\bar{\mathbf{A}}$. Therefore, the rate of convergence in the stochastic dynamics is faster (or slower) than the rate of convergence in the deterministic mean-field approximation if $\sigma_b^2 \geq \sigma_d^2$ (or $\sigma_b^2 < \sigma_d^2$). Note that these inequalities have to be reversed for the rate of divergence from 0 to be faster (or slower) in the stochastic dynamics than that in the mean-field approximation in the case where 0 is SLU with $w(\bar{b} - \bar{d}) > 0$. In particular, the growth rate is always faster in the stochastic dynamics when $\sigma_b^2 = \sigma_d^2$. Analogous conclusions can be drawn for the fixation state $\hat{x} = 1$.

Similarly, in the situation where $\hat{u}(a_t - c_t) = d_t - b_t$ for all $t \geq 0$ with $\hat{u}$ being a positive constant corresponding to a SLS interior equilibrium $\hat{x} = \hat{u}/(1 + \hat{u})$ in the stochastic dynamics under weak selection (that is, $\bar{c} > \bar{a}$), it can be shown that the rate of convergence to $\hat{x}$ in the stochastic dynamics is faster (or slower) than that of in deterministic mean-field approximation if $\hat{u}\big(\sigma_c^2 - \sigma_a^2\big) + 2(\sigma_{c,d} - \sigma_{a,d}) \leq 0$ (or $\hat{u}\big(\sigma_c^2 - \sigma_a^2\big) + 2(\sigma_{c,d} - \sigma_{a,d}) > 0$) (see **Appendix C** for a proof). Moreover, the same is true for the rate of divergence from a SLU $\hat{x}$ in the stochastic dynamics under weak selection (that is, $\bar{c} < \bar{a}$) if the inequalities are reversed. Note that, in the special case where $\sigma_c^2 = \sigma_a^2$ and $\sigma_{c,d} = \sigma_{a,d}$, the growth rate is always faster in the stochastic dynamics.

All these results show that, although stochastic local stability or instability of an equilibrium state may become unaffected by environmental noise as the intensity of selection diminishes, the rate of convergence or divergence of the system near the equilibrium not only depends on environmental noise, but it can be even enhanced by environmental noise. These findings are supported by simulation results presented in **Fig. 3.2**.

3.5 Effect of weak selection on stochastic evolutionary stability

Evolutionary stability, or *evolutionarily stable strategy* (ESS), is the key concept in evolutionary game theory (Maynard Smith, 1982; Lessard, 1984; Hofbauer and Sigmund, 1998). Recently, Zheng *et al.* (2017) extended the standard definition of an ESS in a constant environment (Maynard Smith, 1982) to a variable environment. A *stochastically evolutionarily stable* (SES) strategy is defined as a strategy such that, if all the members of the population adopt it, then the probability for at least any slight perturbed strategy

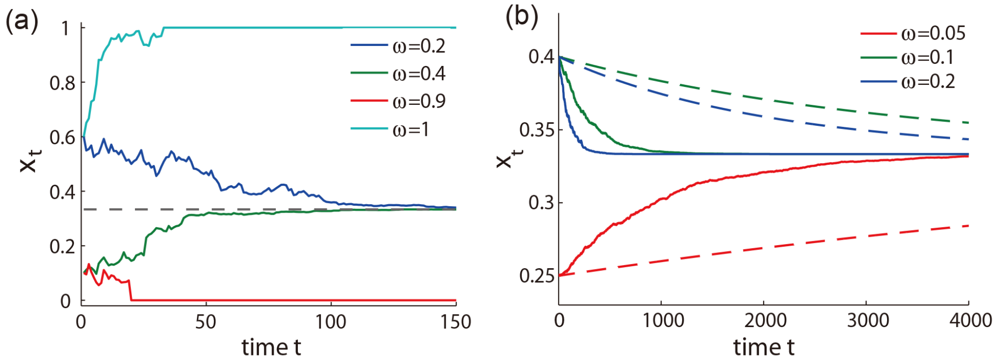


Figure 3.2: *Simulations for the stochastic local stability of a constant interior equilibrium under weak selection.* We consider the random payoff matrix $\begin{pmatrix} 4+z_t & 3 \\ 4 & 3+\hat{u}z_t \end{pmatrix}$, where $\hat{u} = 1/2$ is a positive constant and z_t is taken as a normal random variable with mean $\bar{z} = -0.01$ and variance $\sigma_z^2 = 4$ at time step $t \geq 0$. In this case, $\hat{x} = \hat{u}/(1+\hat{u}) = 1/3$ is a constant interior equilibrium. Moreover: (i) if $w = 1$ (i.e., strong selection), then $\hat{x} = 1/3$ is SLU with respect to the stochastic dynamics; while (ii) $\hat{x} = 1/3$ is a globally asymptotically stable equilibrium with respect to the deterministic dynamics with payoff matrix $\begin{pmatrix} 4+\bar{z} & 3 \\ 4 & 3+\hat{u}\bar{z} \end{pmatrix}$. The simulations show in panel (**a**) that a decrease in the selection intensity w results in $\hat{x} = 1/3$ becoming SLS, and in panel (**b**) that the system state x_t tends to $\hat{x} = 1/3$ when the selection intensity is small enough. Here, each of the solid curves represents an average of 100 simulated curves starting at the same initial state, and the dashed curves represent the deterministic dynamics with payoff matrix given by the mean payoff matrix in the stochastic dynamics.

to invade the population under the influence of natural selection is arbitrarily low. More specifically, a strategy represented by a frequency vector $\hat{\mathbf{x}}$ is SES if $\hat{\mathbf{x}}$-fixation is SLS against any other strategy $\mathbf{x} \neq \hat{\mathbf{x}}$ at least nearby enough (Zheng *et al.*, 2017). Here we mainly focus on the effect of weak selection on stochastic evolutionary stability.

For two mixed strategies $\mathbf{x} = \big(x, 1-x\big)$ and $\hat{\mathbf{x}} = \big(\hat{x}, 1-\hat{x}\big)$ with the payoffs to the pure strategies at time step $t \geq 0$ given by $\mathbf{A}(t)$ in Eq. (3.2.1), the payoff matrix takes the form

$$\begin{pmatrix} \mathbf{x}\mathbf{A}(t)\mathbf{x} & \mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}} \\ \hat{\mathbf{x}}\mathbf{A}(t)\mathbf{x} & \hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}} \end{pmatrix} , \tag{3.5.1}$$

where $\mathbf{x}\mathbf{A}(t)\mathbf{x}$ (respectively, $\mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}}$) is the expected payoff to strategy $\mathbf{x}$ against strategy $\mathbf{x}$ (respectively, $\hat{\mathbf{x}}$), and $\hat{\mathbf{x}}\mathbf{A}(t)\mathbf{x}$ (respectively, $\hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}}$) the expected payoff to strategy $\hat{\mathbf{x}}$ against strategy $\mathbf{x}$ (respectively, $\hat{\mathbf{x}}$). Analogously to the condition Eq. (3.3.7) for the fixation state $\hat{x} = 0$ to be SLS, the fixation of strategy $\hat{\mathbf{x}}$ is SLS if

$$\begin{aligned} &\log\left[\frac{(1-w) + w\left\langle \mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}}\right\rangle}{(1-w) + w\left\langle \hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}}\right\rangle}\right] \\ &+ \frac{w^2\sigma^2_{\hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}}}}{2\big((1-w) + w\left\langle \hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}}\right\rangle\big)^2} - \frac{w^2\sigma^2_{\mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}}}}{2\big((1-w) + w\left\langle \mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}}\right\rangle\big)^2} < 0 , \end{aligned} \tag{3.5.2}$$

where $\sigma^2_{\hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}}}$ and $\sigma^2_{\mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}}}$ denote the variances of $\hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}}$ and $\mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}}$, respectively. Therefore, under weak selection, the fixation of strategy $\hat{\mathbf{x}}$ is SLS if $\big\langle \hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}}\big\rangle - \big\langle \mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}}\big\rangle > 0$, that is, $\hat{\mathbf{x}}\bar{\mathbf{A}}\hat{\mathbf{x}} - \mathbf{x}\bar{\mathbf{A}}\hat{\mathbf{x}} > 0$. Similarly, under weak selection, the fixation of strategy $\mathbf{x}$ is SLU if we have $\big\langle \mathbf{x}\mathbf{A}(t)\mathbf{x}\big\rangle - \big\langle \hat{\mathbf{x}}\mathbf{A}(t)\mathbf{x}\big\rangle < 0$, that is, $\mathbf{x}\bar{\mathbf{A}}\mathbf{x} - \hat{\mathbf{x}}\bar{\mathbf{A}}\mathbf{x} < 0$. Combining these results, we can conclude that, under weak selection, strategy $\hat{\mathbf{x}}$ is SES if and only if

$$\hat{\mathbf{x}}\bar{\mathbf{A}}\hat{\mathbf{x}} - \mathbf{x}\bar{\mathbf{A}}\hat{\mathbf{x}} \ \geq \ 0 \quad \text{for all } \mathbf{x} \neq \hat{\mathbf{x}} , \tag{3.5.3}$$

$$\text{and} \quad \hat{\mathbf{x}}\bar{\mathbf{A}}\mathbf{x} - \mathbf{x}\bar{\mathbf{A}}\mathbf{x} \ > \ 0 \quad \text{for all } \mathbf{x} \neq \hat{\mathbf{x}} \ \text{ if the equality holds above.} \tag{3.5.4}$$

Therefore, under weak selection, a SES strategy is a strategy such that, if all the members of the population adopt it, then the probability for any mutant strategy to invade the population under the influence of natural selection is arbitrarily low.

The above conclusion shows that the conditions for strategy $\hat{\mathbf{x}}$ to be SES under weak selection depends only on the average payoff matrix $\bar{\mathbf{A}}$, and that they exactly match the standard conditions for an ESS with the payoff matrix $\bar{\mathbf{A}}$ (Maynard Smith, 1982). So, under weak selection: (i) the pure strategy $\hat{\mathbf{x}} = (0, 1)$ is SES if $\bar{d} > \bar{b}$; (ii) the pure strategy $\hat{\mathbf{x}} = (1, 0)$ is SES if $\bar{a} > \bar{c}$; and (iii) if $\bar{a} > \bar{c}$ and $\bar{d} > \bar{b}$, or $\bar{a} < \bar{c}$ and $\bar{d} < \bar{b}$, then the mixed strategy $\hat{\mathbf{x}} = (\hat{x}, 1 - \hat{x})$ with $\hat{x} = (\bar{b} - \bar{d})/(\bar{b} - \bar{d} + \bar{c} - \bar{a})$ is SES if $\bar{b} > \bar{d}$ and $\bar{c} > \bar{a}$ (Maynard Smith, 1982; Hofbuaer and Sigmund, 1998). Moreover, even if no constant interior equilibrium exists in Eq. (3.2.2), it is still possible for a mixed strategy to be SES. For example, consider a random payoff matrix $\begin{pmatrix} 1 + \xi_t & 3 \\ 3 & 2 + \xi_t \end{pmatrix}$, where ξ_t is a random variable with mean $\langle \xi_t \rangle = 0$ and variance $\langle \xi_t^2 \rangle = \sigma_\xi^2$ at time step $t \geq 0$, where σ_ξ^2 is small but $\sigma_\xi^2 \neq 0$ such that both $1 + \xi_t$ and $2 + \xi_t$ are positive random payoffs for $t \geq 0$. With this random payoff matrix, although no constant interior equilibrium exists, the mixed strategy $\hat{\mathbf{x}} = (\hat{x}, 1 - \hat{x})$ with $\hat{x} = 1/3$ is SES with respect to the stochastic dynamics.

3.6 Discussion

How natural selection can reduce the impact of environmental stochastic fluctuations on the evolution of animal behavior is a very challenging question. In this study, we have considered the effects of weak selection on a two-phenotype evolutionary game dynamics in an infinite population with a random payoff matrix. The results show that, under weak selection, both stochastic local stability and stochastic evolutionary stability in this system depend only on the means of the random payoffs and not at all on their variances. However, although stochastic local stability or instability of an equilibrium may not be affected by environmental noise, the rate of convergence or divergence near an equilibrium not only depends on environmental noise, but it can even be enhanced by environmental noise. This is the case, for instance, when the variances of the random payoffs are equal as well as the covariances. These predictions are supported by analytical approximations and computer simulations.

Our analysis is based on the concept of stochastic evolutionary stability (SES) through the analysis of stochastic local stability (SLS) that was developed in a previous paper of ours (Zheng *et al.*, 2017) to predict the results of long-term evolution of strategies in a stochastic environment. This is actually an extension of the classic concept of an evolutionarily stable strategy (ESS) to take into account random payoffs as a result of environmental noise. These have been approximated in the case of weak selection to show that stochastic evolutionary stability can be unaffected, and evolution can even occur faster, in the presence of environmental noise when selection is weak enough. It might be worth stressing that weak selection is not equivalent to weak noise. Actually, it is almost the opposite, since selection would appear often strong when noise is weak. It may be obvious that the effects of weak noise can be counteracted by the pressure of strong selection. That the effects of noise can be counteracted by the pressure of weak selection is less obvious, not to mention that weak selection can increase the rate of evolution in the presence of noise. These findings have biological implications, since they reveal an unexpected role of weak selection in the evolution of biological populations in a random environment.

Previous studies on the impact of environmental noise on biological evolution involved such mechanisms as the storage effect and the bet-hedging strategy in populations with overlapping generations (Chesson, 1983; Warner and Chesson, 1985; Ellner and Hairston, 1994; Olofsson *et al.*, 2009). Such mechanisms concern the trade-off between adult survival and reproduction, but can involve, in principle, any life history trait. They have been used to explain the coexistence of competitors, and are somehow related to the notion of protected polymorphism in structured populations under the effects of spatially or temporally varying selection regimes (Karlin and Liberman, 1982). Our study takes the opposite view of looking at a general condition, namely, weak selection, which could counteract the effects of random noise. That the same condition can enhance the rate of evolution in the presence of random noise is an unexpected bonus. And that the results are obtained under minimal assumptions, namely a matrix game with random payoffs in a well-mixed population, suggest that they might be of general validity.

3.7 Appendix

3.7.1 A. Stochastic local stability of fixation state $\hat{x} = 0$ in degenerate cases

In the degenerate case where $b_t = d_t$ for all $t \geq 0$, let $v_t = 1/u_t = (1 - x_t)/x_t$. From Eq. (3.2.3), we have the recurrence equation

$$v_{t+1} = v_t \left[\frac{\big((1-w) + wc_t\big) + v_t\big((1-w) + wd_t\big)}{\big((1-w) + wa_t\big) + v_t\big((1-w) + wd_t\big)} \right] . \tag{3.7.1}$$

Iterating this recurrence equation leads to

$$\begin{aligned} \frac{1}{n}\big(v_n - v_0\big) &= \frac{1}{n} \sum_{t=0}^{n-1} \left[\frac{(1-w) + wc_t}{(1-w) + wd_t} - \frac{(1-w) + wa_t}{(1-w) + wd_t} \right] \\ &\quad - \frac{1}{n} \sum_{t=0}^{n-1} \frac{\frac{(1-w)+wa_t}{(1-w)+wd_t} \left(1 - \frac{(1-w)+wa_t}{(1-w)+wc_t}\right)}{\frac{(1-w)+wa_t}{(1-w)+wc_t} + \frac{(1-w)+wd_t}{(1-w)+wc_t} v_t} . \end{aligned} \tag{3.7.2}$$

Therefore, if $u_t \to 0$ (that is, $v_t \to \infty$), then the strong law of large numbers guarantees that

$$\lim_{n \to \infty} \frac{1}{n}\big(v_n - v_0\big) \approx \left\langle \frac{(1-w) + wc_t}{(1-w) + wd_t} - \frac{(1-w) + wa_t}{(1-w) + wd_t} \right\rangle . \tag{3.7.3}$$

Then, using Egorov's theorem, the fixation state $\hat{x} = 0$ is SLS if

$$\left\langle \frac{(1-w) + wc_t}{(1-w) + wd_t} \right\rangle - \left\langle \frac{(1-w) + wa_t}{(1-w) + wd_t} \right\rangle > 0 \tag{3.7.4}$$

(the more rigorous mathematical proofs are similar to those in Zheng *et al.* (2017)).

Note that

$$\left\langle \frac{(1-w) + wc_t}{(1-w) + wd_t} \right\rangle \approx \frac{(1-w) + w\bar{c}}{(1-w) + w\bar{d}} + \frac{\big((1-w) + w\bar{c}\big)w^2\sigma_d^2}{\big((1-w) + w\bar{d}\big)^3} - \frac{w^2\sigma_{c,d}^2}{\big((1-w) + w\bar{d}\big)^2} ,$$

and

$$\left\langle \frac{(1-w)+wa_t}{(1-w)+wd_t} \right\rangle \approx \frac{(1-w)+w\bar{a}}{(1-w)+w\bar{d}} + \frac{\big((1-w)+w\bar{a}\big)w^2\sigma_d^2}{\big((1-w)+w\bar{d}\big)^3} - \frac{w^2\sigma_{a,d}^2}{\big((1-w)+w\bar{d}\big)^2} \ .$$

Thus, under weak enough selection (that is, for w small enough), the fixation state $\hat{x} = 0$ is SLS if $\bar{c} - \bar{a} > 0$, and SLU if $\bar{c} - \bar{a} < 0$.

Similarly, in the degenerate case where $a_t = c_t$ for all $t \geq 0$, under weak enough selection, the fixation state $\hat{x} = 1$ is SLS if $\bar{b} - \bar{d} > 0$, and SLU if $\bar{b} - \bar{d} < 0$.

3.7.2 B. Stochastic local stability of a constant interior equilibrium

With the random payoff matrix $\mathbf{A}(t)$ in Eq. (3.3.9) where $\hat{u} > 0$, the recurrence equation in Eq. (3.2.3) can be rewritten in the form

$$u_{t+1} = u_t \left[\frac{u_t[(1-w)+w(c_t+z_t)] + [(1-w)+wb_t]}{u_t[(1-w)+wc_t] + [(1-w)+w(b_t+\hat{u}z_t)]} \right] \ . \tag{3.7.5}$$

From this equation and the equality $\hat{u}(a_t - c_t) = d_t - b_t$, we have

$$\begin{aligned} u_{t+1} - \hat{u} &= \big(u_t - \hat{u}\big) \left[\frac{u_t((1-w)+wc_t) + u_t w z_t + \hat{u} w z_t + ((1-w)+wb_t)}{u_t((1-w)+wc_t) + ((1-w)+wb_t) + \hat{u} w z_t} \right] \\ &= \big(u_t - \hat{u}\big) \left[\frac{u_t((1-w)+wa_t) + ((1-w)+wd_t)}{u_t((1-w)+wc_t) + ((1-w)+wd_t)} \right] \ . \end{aligned} \tag{3.7.6}$$

In particular, this ensures that $u_{t+1} - \hat{u} > 0$ if $u_t - \hat{u} > 0$, and $u_{t+1} - \hat{u} < 0$ if $u_t - \hat{u} < 0$. Moreover, some algebraic manipulations yield

$$\frac{u_t((1-w)+wa_t) + ((1-w)+wd_t)}{u_t((1-w)+wc_t) + ((1-w)+wd_t)} = \left[\frac{\hat{u}((1-w)+wa_t) + ((1-w)+wd_t)}{\hat{u}((1-w)+wc_t) + ((1-w)+wd_t)} \right] Q_t \ ,$$

where

$$Q_t = 1 - \frac{(u_t - \hat{u})((1-w)+wd_t)wz_t}{D_t} \tag{3.7.7}$$

with

$$\begin{aligned} D_t = &\Big[\hat{u}((1-w)+wa_t) + ((1-w)+wd_t)\Big] \\ &\times \Big[\hat{u}((1-w)+wc_t) + ((1-w)+wd_t) + (u_t - \hat{u})((1-w)+wc_t)\Big] \ . \end{aligned} \tag{3.7.8}$$

Therefore, iterating Eq. (3.7.6) leads to

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\log\left[\frac{u_n-\hat{u}}{u_0-\hat{u}}\right] &= \lim_{n\to\infty}\frac{1}{n}\sum_{t=0}^{n-1}\log\left[\frac{\hat{u}((1-w)+wa_t)+((1-w)+wd_t)}{\hat{u}((1-w)+wc_t)+((1-w)+wd_t)}\right] \\ &\quad + \lim_{n\to\infty}\frac{1}{n}\sum_{t=0}^{n-1}\log Q_t \ . \end{aligned} \tag{3.7.9}$$

If $u_t \to \hat{u}$ (which compels $Q_t \to 1$), then the strong law of large numbers guarantees that

$$\lim_{n\to\infty}\frac{1}{n}\log\left[\frac{u_n-\hat{u}}{u_0-\hat{u}}\right] \approx \left\langle \log\left[\frac{\hat{u}((1-w)+wa_t)+((1-w)+wd_t)}{\hat{u}((1-w)+wc_t)+((1-w)+wd_t)}\right]\right\rangle. \tag{3.7.10}$$

Using Egorov's theorem, the constant interior equilibrium $\hat{x} = \hat{u}/(1+\hat{u})$ is SLS if

$$\left\langle \log\left[\frac{\hat{u}((1-w)+wa_t)+((1-w)+wd_t)}{\hat{u}((1-w)+wc_t)+((1-w)+wd_t)}\right]\right\rangle < 0 \ , \tag{3.7.11}$$

and SLU if the inequality is reversed.

Note that

$$\begin{aligned}\left\langle \log\Big[\hat{u}((1-w)+wa_t)+((1-w)+wd_t)\Big]\right\rangle &\approx \log\Big[\hat{u}((1-w)+w\bar{a})+((1-w)+w\bar{d})\Big] \\ &\quad - \frac{\hat{u}^2w^2\sigma_a^2+w^2\sigma_d^2+2\hat{u}w^2\sigma_{a,d}}{2\big[\hat{u}((1-w)+w\bar{a})+((1-w)+w\bar{d}(\big]^2}\end{aligned}$$

and

$$\begin{aligned}\left\langle \log\Big[\hat{u}((1-w)+wc_t)+((1-w)+wd_t)\Big]\right\rangle &\approx \log\Big[\hat{u}((1-w)+w\bar{c})+((1-w)+w\bar{d})\Big] \\ &\quad - \frac{\hat{u}^2w^2\sigma_c^2+w^2\sigma_d^2+2\hat{u}w^2\sigma_{c,d}}{2\big[\hat{u}((1-w)+w\bar{c})+((1-w)+w\bar{d})\big]^2} \ .\end{aligned}$$

Thus, under weak enough selection, $\hat{x} = \hat{u}/(1+\hat{u})$ is SLS if $\bar{c}-\bar{a} > 0$, and SLU if $\bar{c}-\bar{a} < 0$.

3.7.3 C. Convergence rate near a SLS constant interior equilibrium

With the random payoff matrix $\mathbf{A}(t)$ in Eq. (3.3.9), we have shown that, under weak enough selection, the constant interior equilibrium $\hat{x} = \hat{u}/(1+\hat{u})$ is SLS if $\bar{c}-\bar{a} > 0$. When the system state is near this constant interior equilibrium, the convergence rate of the system to it is given by the right-hand member in Eq. (3.7.10). Under weak selection, the convergence rate is approximated as

$$\begin{aligned} &\left\langle \log\left[\frac{\hat{u}((1-w)+wa_t)+((1-w)+wd_t)}{\hat{u}((1-w)+wc_t)+((1-w)+wd_t)}\right]\right\rangle \\ &\approx \log\left[\frac{\hat{u}((1-w)+w\bar{a})+((1-w)+w\bar{d})}{\hat{u}((1-w)+w\bar{c})+((1-w)+w\bar{d})}\right] \\ &-\frac{w^2\big(\hat{u}^2\sigma_a^2+\sigma_d^2+2\hat{u}\sigma_{a,d}\big)}{2\big[\hat{u}((1-w)+w\bar{a})+((1-w)+w\bar{d})\big]^2} \\ &+\frac{w^2\big(\hat{u}^2\sigma_c^2+\sigma_d^2+2\hat{u}\sigma_{c,d}\big)}{2\big[\hat{u}((1-w)+w\bar{c})+((1-w)+w\bar{d})\big]^2}\ , \end{aligned} \tag{3.7.12}$$

where the term

$$\log\left[\frac{\hat{u}((1-w)+w\bar{a})+((1-w)+w\bar{d})}{\hat{u}((1-w)+w\bar{c})+((1-w)+w\bar{d})}\right]$$

corresponds to the convergence rate of the deterministic system with payoff matrix $\bar{\mathbf{A}}$. Furthermore, if w is small enough, we have the approximations

$$\log\left[\frac{\hat{u}((1-w)+w\bar{a})+((1-w)+w\bar{d})}{\hat{u}((1-w)+w\bar{c})+((1-w)+w\bar{d})}\right] \approx \frac{w(\bar{a}-\bar{c})}{(1+\hat{u})}\ , \tag{3.7.13}$$

$$\frac{w^2\big(\hat{u}^2\sigma_a^2+\sigma_d^2+2\hat{u}\sigma_{a,d}\big)}{2\big[\hat{u}((1-w)+w\bar{a})+((1-w)+w\bar{d})\big]^2} \approx \frac{w^2\big(\hat{u}^2\sigma_a^2+\sigma_d^2+2\hat{u}\sigma_{a,d}\big)}{2(1+\hat{u})^2}\ , \tag{3.7.14}$$

$$\frac{w^2\big(\hat{u}^2\sigma_c^2+\sigma_d^2+2\hat{u}\sigma_{c,d}\big)}{2\big[\hat{u}((1-w)+w\bar{c})+((1-w)+w\bar{d})\big]^2} \approx \frac{w^2\big(\hat{u}^2\sigma_c^2+\sigma_d^2+2\hat{u}\sigma_{c,d}\big)}{2(1+\hat{u})^2}\ . \tag{3.7.15}$$

Therefore, Eq. (3.7.12) can be rewritten as

$$\begin{aligned} &\left\langle \log\left[\frac{\hat{u}((1-w)+wa_t)+((1-w)+wd_t)}{\hat{u}((1-w)+wc_t)+((1-w)+wd_t)}\right]\right\rangle \approx \frac{w(\bar{a}-\bar{c})}{(1+\hat{u})} \\ &+\begin{cases} \frac{\hat{u}w^2}{2(1+\hat{u})^2}\big[\hat{u}\big(\sigma_c^2-\sigma_a^2\big)+2(\sigma_{c,d}-\sigma_{a,d})\big] & \text{if } \ \hat{u}\big(\sigma_c^2-\sigma_a^2\big)+2(\sigma_{c,d}-\sigma_{a,d})\neq 0\ , \\ \frac{w^3}{(1+\hat{u})^3}\big(\hat{u}^2\sigma_a^2+\sigma_d^2+2\hat{u}\sigma_{a,d}\big)(\bar{a}-\bar{c}) & \text{if } \ \hat{u}\big(\sigma_c^2-\sigma_a^2\big)+2(\sigma_{c,d}-\sigma_{a,d})=0\ . \end{cases} \end{aligned} \tag{3.7.16}$$

This implies that, under weak enough selection, the convergence rate near the SLS constant interior equilibrium $\hat{u}>0$ in the stochastic dynamics (with $\bar{c}>\bar{a}$) if faster (or slower) than that in the deterministic mean field approximation if $\hat{u}\big(\sigma_c^2-\sigma_a^2\big)+2(\sigma_{c,d}-\sigma_{a,d})\leq 0$ (or $\hat{u}\big(\sigma_c^2-\sigma_a^2\big)+2(\sigma_{c,d}-\sigma_{a,d})>0$).

Acknowledgements

In this study, X.-D.Z. and Y.T. were supported by the National Natural Science Foundation of China (Grants No.31971511 and No.31770426,) and the Ministry of Science and Technology of China (Grant No.2018YFC1003300); C.L. and S.L. were supported in part by Natural Sciences and Engineering Research Council of Canada, Grant No.8833; S.L. was supported also by Chinese Academy of Sciences President's International Fellowship Initiative (Grant No.2016VBA039).

C.L. and X.-D.Z. contributed equally to this work.

Bibliography

[1] M. Kimura, Evolutionary rate at the molecular level. *Nature* **217**, 624-626 (1968).

[2] T. Ohta, Near-neutrality in evolution of genes and gene regulation. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 16134-16137 (2002).

[3] J. Hofbauer, K. Sigmund, *The Theory of Evolution and Dynamical Systems* (Cambridge University Press, Cambridge, England, 1998).

[4] A. Traulsen, J.C. Claussen, C. Hauert, Coevolutionary dynamics: From finite to infinite populations. *Physical Review Letters* **95**, 238701 (2005).

[5] M.A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646-650 (2004).

[6] H. Ohtsuki, C. Hauert, E. Lieberman, M. A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502-505 (2006).

[7] M. A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).

[8] C. Taylor, D. Fudenberg, A. Sasaki, M. A. Nowak, Evolutionary game dynamics in finite populations. *Bulletin of Mathematical Biology* **66**, 1621 (2004).

[9] S. Lessard, Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory. *Theoretical Population Biology* **68**, 19-27 (2005).

[10] A. Traulsen, M.A. Nowak, J.M. Pacheco, Stochastic dynamics of invasion and fixation. *Physical Review E* **74**, 011909 (2006a).

[11] A. Traulsen, J.M. Pacheco, L.A. Imhof, Stochasticity and evolutionary stability . *Physical Review E* **74**, 021905 (2006b).

[12] S. Lessard, V. Ladret, The probability of fixation of a single mutant in an exchangeable selection model. *Journal of Mathematical Biology* **54**, 721-744 (2007).

[13] A. Traulsen, C. Hauert, Stochastic evolutionary game dynamics. *Reviews of Nonlinear Dynamics and Complexity*, **Vol. II** (Wiley-VCH, edited by H.G. Schuster, 2009)

[14] X.-D. Zheng, R. Cressman, Y. Tao, The diffusion approximation of stochastic evolutionary game dynamics: mean effective fixation time and the significance of the one-third law. *Dynamic Games and Applications* **1**, 462-477 (2011).

[15] B. Wu, P.M. Altrock, L. Wang, A. Traulsen, Universality of weak selection. *Physical Review E* **82**, 046106 (2010).

[16] B. Wu, J. Garcia, C. Hauert, A. Traulsen, Extrapolating weak selection in evolutionary games. *PLoS Computational Biology* **9**, e1003381 (2013).

[17] X.-D. Zheng, C. Li, S. Lessard, Y. Tao, Evolutionary stability concepts in a stochastic environment. *Physical Review E* **96**, 032414 (2017).

[18] X.-D. Zheng, C. Li, S. Lessard, Y. Tao, Environmental noise could promote stochastic local stability of behavioral diversity evolution. *Physical Review Letters* **120**, 218101 (2018).

[19] R.M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, 1973)

[20] F. Stollmeier, J. Nagler, Unfair and anomalous evolutionary dynamics from fluctuating payoffs. *Physical Review Letters* **120**, 058101 (2018).

[21] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, England, 1982).

[22] S. Lessard, Evolutionary dynamics in frequency-dependent two-phenotype models. *Theoretical Population Biology* **25**, 210-234 (1984).

[23] S. Karlin, U. Liberman, Random temporal variation in selection intensities: case of large population size. *Theoretical Population Biology* **6**, 355-382 (1974).

[24] S. Karlin, U. Liberman, Random temporal variation in selection intensities: One-locus two-allele model. *Journal of Mathematical Biology* **2**, 1-17 (1975).

[25] P.L. Chesson, Coexistence of competitors in a stochastic environment:The storage effect. *Lecture Notes Biomath* **52**, 188-198 (1983).

[26] R.R. Warner, P.L. Chesson, Coexistence mediated by recruitment fluctuations: a field guide to the storage effect. *American Naturalist* **125**, 769-787 (1985).

[27] H. Olofsson, J. Ripa, N. Jonzén, Bet-hedging as an evolutionary game: the trade-off between egg size and number. *Proceedings of the Royal Society B* **276**, 2963-2969 (2009).

[28] S. Ellner, N.G. Hairston, Role of overlapping generations in maintaining genetic variation in a fluctuating environment. *American Naturalist* **143**, 403-417 (1994).

[29] S. Karlin, Classifications of selection-migration structures and conditions for a protected polymorphism, in *Evolutionary Biology*, vol. 14, 61 (Plenum Publising Corporation, New York, edited by M.K. Hecht, B. Wallace and C.T. Prance, 1982).

Chapter 4

Article 4

Uncertainty in the payoffs for defection could be conducive to the evolution of cooperation

Cong Li[1,2,*], Ting Ji[1,*], Qiao-Qiao He[3,*], Xiu-Deng Zheng[1], Sabin Lessard[2] and Yi Tao[1,4]

[1]Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

[2]Department of Mathematics and Statistics, University of Montreal, Montreal, Quebec H3C 3J7, Canada

[3]College of Life Science, Shenyang Normal University, Shenyang 110034, China

[4]University of Chinese Academy of Sciences, Beijing 100049, China

[*]These authors contributed equally to this work.

Information

This paper is currently under review. It was submitted to *Journal of Theoretical Biology* on 27 March 2020.

My contribution

I did most of the theoretical analysis and revised the paper based on a first draft by Yi Tao. I worked on all the examples and ran the numerical simulations.

Abstract

Stochastic variation in the environment is one of the most important characteristics of nature, and it may cause changes in the occurrence of interactions between individuals and, more importantly, in the payoffs received by interacting individuals. In order to explore the effects of random environmental noise on the evolution of cooperation, the classic Prisoner's Dilemma (PD) game is extended to a more general randomized Prisoner's Dilemma (RPD) game. The entries of the payoff matrix in a RPD game are random variables whose average values determine a classic PD game. The concepts of stochastic local stability (SLS) and stochastic evolutionary stability (SES) applied to the RPD game suggest that the evolution of cooperation could be made easier if the coefficients of variation of the payoffs are smaller for the cooperation strategy than for the defection strategy.

4.1 Introduction

In 1950, based on experiments made by Merril Flood and Melvin Dresher, Albert Tucker proposed a simple but illuminating situation in order to model cooperation, the Prisoner's Dilemma (PD) (Poundstone, 1992; Nowak and Highfield, 2011). Imagine that you and your accomplice are both held prisoners, having been charged with a serious crime. The prosecutor offers each one of you the following deal: if you plead guilty, while your partner remains silent, then you will be sentenced to 1 year, and your partner to 4 years; if you and your partner remain both silent, then you will each be sentenced to 2 years; and if you and your partner plead both guilty, then you will each be sentenced to 3 years. In this situation, plead guilty or remain silent means defection or cooperation with the partner, respectively. The rational choice, whatever the partner does, is to plead guilty. Defection by both players is the only Nash equilibrium (NE) in this two-player game (Hofbauer and Sigmund, 1998; Nowak, 2006).

In the previous situation, however, if only the court (not the prosecutor) has the right to determine the final sentence, and if the deal that the prosecutor can offer is expressed only in general terms (for example, be frank and expect a lenient sentence, or resist and expect a severe sentence), then what would be the rational choice? For illustration, consider the situation where the prosecutor offers the following deal to each one of the two accomplices: if you plead guilty, while your partner remains silent, then expect to be released or sentenced to 1 or 2 years with equal probability 1/3, while your partner will be sentenced to 3, 4, or 5 years with equal probability 1/3; if you and your partner remain both silent, then expect each of you to be sentenced to 1, 2 or 3 years with equal probability 1/3; and if you and your partner plead both guilty, then expect each of you to be sentenced to 2, 3 or 4 years with equal probability 1/3. What distinguishes this situation from the previous one is the uncertainty (or randomness) in the sentences since the expected values are the same. A natural question pertains to the consequences of non-deterministic outcomes on rational choices. Note that these have to be distinguished from consequences of previous outcomes on available actions and decision making which lead to what is known as stochastic games (Fundenberg *et al.*, 2012; Solan an Vieille, 2015; Hilbe *et al.*, 2018).

Since the 1950s, the PD game has been used as a theoretical framework to explain the evolution of cooperation in natural populations and human societies (Axelrod and Hamilton 1981; Maynard Smith, 1982; Axelrod, 1984; Hofbauer and Sigmund, 1998; Nowak, 2006). In the standard PD game, only two phenotypes (or strategies) are considered, cooperation (C) and defection (D), and the payoff matrix has entries R, S, T, P with $T > R > P > S$, where R and S are the constant payoffs to C against C and D, respectively, while T and P are the corresponding payoffs to D against the same two strategies. In this game, defection is the only NE. In order to explain how cooperation could be favored by natural selection in the framework of the PD game, several mechanisms have been proposed (Nowak, 2006; Nowak and Sigmund, 2007), among which and not the least direct reciprocity (Trivers, 1971; Axelrod, 1984). However, this mechanism and all others are usually considered under the assumption that the payoff matrix of the underlying PD game is a constant matrix.

Environmental conditions in the real world are obviously changing and uncertain, and these may cause changes in the occurrence of interactions between individuals and, more importantly, in the payoffs received by interacting individuals. Therefore, unless stochastic fluctuations are small enough so that their effects can be neglected, there is no *a priori* reason to assume that the payoff matrix of an evolutionary game is constant if the environment is actually stochastic. In fact, effects of environmental stochasticity on population and community ecology have been investigated by many authors (Lande *et al.*, 2003). May (1973), for instance, pointed out that birth rates, carrying capacities, competition coefficients, and other parameters which characterize natural biological systems, to a greater or lesser degree, exhibit random fluctuations.

Recently, in order to study environmental stochastic effects on the evolution of animal behavior, we developed the concepts of stochastic evolutionary stability (SES) and stochastic convergence stability (SCS) by investigating conditions for stochastic local stability (Zheng *et al.*, 2017, 2018). Following Karlin and Lieberman (Karlin and Liberman, 1974, 1975; Ewens, 2004), a constant equilibrium is said to be *stochastically locally stable* (SLS) if the population state tends to this equilibrium as time goes to infinity with probability arbitrarily close to 1 when the initial population state is sufficiently near the equilibrium (Zheng *et al.*, 2017). On the contrary, if the population state cannot tend to the equilibrium from any other state with probability one, then the equilibrium is said to be *stochastically locally unstable* (SLU). On the other hand, extending the standard definition of an evolutionarily stable strategy (ESS) (Maynard Smith and Price, 1973; Maynard Smith, 1974, 1982) to a variable environment, a strategy is said to be *stochastically evolutionarily stable* (SES) if, when all members of the population adopt it, then the probability for at least any slight perturbed strategy to invade the population under the effects of natural selection is arbitrarily small (Zheng *et al.*, 2017). However, in the context of the PD game, a natural question to ask is how the evolution of cooperation is influenced by random fluctuations in the payoffs as a result of environmental noise.

In fact, some studies have involved the effect of variability of payoffs on the evolution of cooperation. For example, Johnson *et al.* (2002) considered whether the stochasticity of payoffs in a PD game will affect the evolution of cooperation, and they emphasized that relaxing the assumption of fixed payoffs leads to frequent violations of the payoff structure required for a classic PD game; and Perc (2006) also investigated a PD game on a lattices network using simulations, and he found that the increase of stochastic fluctuation intensity in payoffs might be conducive to the evolution of cooperation. Moreover, Szolnoki *et al.* (2019) studied the evolution of cooperation with periodic payoffs on both square lattices and regular random graph.

In this study, based on the concepts of stochastic local stability and stochastic evolutionary stability in matrix games with random payoffs (Zheng *et al.*, 2017), the randomized

Prisoner's Dilemma (RPD) game, where the entries of the payoff matrix are random variables whose average values determine a classic PD game, provides a general theoretical framework for understanding the evolution of cooperation in a stochastic environment. Moreover, we show also mathematically why uncertainty in the payoffs for defection could be conducive to the evolution of cooperation.

4.2 Model and analysis

We consider a PD game in discrete time with a random payoff matrix $\Pi(t)$ at time step $t \geq 0$ whose entries are R_t, S_t, T_t and P_t, where R_t and S_t are the payoffs to C against C and D, respectively, while T_t and P_t are the corresponding payoffs to D against the same two strategies. These payoffs are assumed to be positive random variables that are uniformly bounded below and above by some positive constants (Zheng *et al.*, 2017). Moreover, (R_t, S_t, T_t, P_t) for all integers $t \geq 0$ are assumed to be independent identically distributed random vectors, with means, variances and covariances of the components denoted by $\langle \alpha_t \rangle = \bar{\alpha}$, $\big\langle (\alpha_t - \bar{\alpha})^2 \big\rangle = \sigma_\alpha^2$ and $\big\langle (\alpha_t - \bar{\alpha})(\beta_t - \bar{\beta}) \big\rangle = \sigma_{\alpha\beta}$, respectively, for $\alpha_t, \beta_t = R_t, S_t, T_t, P_t$; and covariances of α_s and β_t equal to 0 for $s \neq t$. It is assumed throughout that the variances and covariances of the random payoffs R_t, S_t, T_t and P_t are small. Furthermore, their means satisfy the inequalities $\bar{T} > \bar{R} > \bar{P} > \bar{S}$. A PD game with a random payoff matrix $\Pi(t)$ for $t \geq 0$ as above describes a randomized Prisoner's Dilemma (RPD).

Let x_t denote the frequency of C at time step $t \geq 0$. Assuming an infinite population with random pairwise interactions at each time step and average payoff as fitness, the frequency of C from time step t to time step $t+1$ is given by the stochastic recurrence equation

$$x_{t+1} \quad = \quad \frac{x_t\big(x_t R_t + (1-x_t)S_t\big)}{x_t\big(x_t R_t + (1-x_t)S_t\big) + (1-x_t)\big(x_t T_t + (1-x_t)P_t\big)} \;, \tag{4.2.1}$$

where $x_t R_t + (1-x_t)S_t$ is the average payoff to C and, similarly, $x_t T_t + (1-x_t)P_t$ the average payoff to D, both at time step $t \geq 0$ (Taylor and Jonker, 1978; Maynard Smith, 1982; Lessard, 1984). The stochastic dynamical properties of this two-phenotype game model with random payoffs are highlighted by a local stability analysis of the fixation states (Zheng *et al.*, 2017) (see also the mathematical analysis in **Appendix A**). It can be shown that C-fixation (i.e., the boundary $x = 1$) is SLS if $\log\left(\bar{R}/\bar{T}\right) > \left(\sigma_R^2/\bar{R}^2 - \sigma_T^2/\bar{T}^2\right)/2$ and SLU if the inequality is reversed. By symmetry, D-fixation (i.e., the boundary $x = 0$) is SLS if $\log\left(\bar{P}/\bar{S}\right) > \left(\sigma_P^2/\bar{P}^2 - \sigma_S^2/\bar{S}^2\right)/2$ and SLU if the inequality is reversed. These conditions imply that the evolution of cooperation could be more favored by natural selection if the coefficients of variation of the payoffs for the cooperation strategy, namely $\sigma_R/\bar{R}$ and $\sigma_S/\bar{S}$, are smaller than those for the defection strategy, namely $\sigma_T/\bar{T}$ and $\sigma_P/\bar{P}$.

In particular, if $\sigma_R^2 = 0$, then C-fixation is SLS if $\sigma_T^2 > -2\bar{T}^2 \log\left(\bar{R}/\bar{T}\right)$ (which implies that an increase of σ_T^2 promotes stochastic local stability of C-fixation (Fig. 4.1A); and if $\sigma_S^2 = 0$, then D-fixation is SLU if $\sigma_P^2 > 2\bar{P}^2 \log\left(\bar{P}/\bar{S}\right)$ (which implies that an increase of σ_P^2 may result in a loss of stochastic local stability of D-fixation) (Fig. 4.1B). Therefore, under the RPD game dynamics, not only C-fixation can be SLS even if $\bar{T} > \bar{R}$, but also D-fixation can be SLU even if $\bar{P} > \bar{S}$. This is in strong contrast with the dynamics in the case of a constant PD game. More importantly, uncertainty (or randomness) in the strategy payoffs could favor the evolution of cooperation.

On the other hand, it can be shown that the pure strategy C is SES against any nearby mixed strategy if $\bar{R}^2(\bar{T} - \bar{R}) > -(\bar{T}\sigma_R^2 - \bar{R}\sigma_{RT})$ (Zheng *et al.*, 2017) (see also the mathematical analysis in **Appendix B**). By symmetry, the pure strategy D is SES against any nearby mixed strategy if $\bar{P}^2(\bar{S} - \bar{P}) < -(\bar{S}\sigma_P^2 - \bar{P}\sigma_{SP})$. Comparing the conditions for stochastic local stability and stochastic evolutionary stability, we find that: (i) if $\sigma_R^2 = 0$, then the pure strategy C may not be SES even in the case where C-fixation is SLS; and (ii) if $\sigma_S^2 = 0$, then the pure strategy D is SES in the case where $\sigma_P^2 < \bar{P}^2(\bar{P} - \bar{S})/\bar{S}$, while D-fixation is SLS in the case where $\sigma_P^2 < 2\bar{P}^2 \log\left(\bar{P}/\bar{S}\right)$. Consequently, there exists a critical value of the ratio $\bar{P}/\bar{S}$, denoted by z^*, such that D-fixation is SLS if $\bar{P}/\bar{S} < z^*$, while the pure strategy D may not be SES, which occurs when $\bar{P}^2(\bar{P}-\bar{S})/\bar{S} < \sigma_P^2 < 2\bar{P}^2 \log\left(\bar{P}/\bar{S}\right)$; and, conversely, D-fixation is SLU if $\bar{P}/\bar{S} > z^*$, while the pure strategy D may be SES, which occurs when $2\bar{P}^2 \log\left(\bar{P}/\bar{S}\right) < \sigma_P^2 < \bar{P}^2(\bar{P}-\bar{S})/\bar{S}$ (see Fig. 4.1C). Thus, in the case of a RPD game, the stochastic local stability of C-fixation (or D-fixation) and the stochastic evolutionary stability of C (or D) could be inconsistent. These dynamical properties of the randomized PD game suggest that cooperation may be more likely to be maintained by natural selection in a stochastic environment, even though it may not be stochastically evolutionarily stable.

In order to test the above theoretical results and predictions, four examples are investigated using computer simulations. In these examples, we consider the probability distribution of the frequency of strategy C after 10^4 time steps over 10^3 runs based on Eq. (4.2.1) starting with an initial frequency $x_0 = 0.5$. The means of the random payoffs are always the same and given by $\bar{R} = 12$, $\bar{S} = 9$, $\bar{T} = 13$ and $\bar{P} = 10$, respectively. For the variances, four sets of values are considered below.

Example 1. $\sigma_R^2 = \sigma_S^2 = \sigma_T^2 = 0$ and $\sigma_P^2 = 90.25/3$ (Fig. 4.2A): C-fixation and D-fixation are both SLU and the probability of $x \leq 0.1$ is only 0.248.

Example 2. $\sigma_R^2 = \sigma_S^2 = 0$, $\sigma_T^2 = 56.25/3$ and $\sigma_P^2 = 90.25/3$ (Fig. 4.2B): C-fixation and D-fixation are both SLU and the probability of $x \leq 0.1$ is only 0.175.

Example 3. $\sigma_R^2 = \sigma_S^2 = 0$, $\sigma_T^2 = 100/3$ and $\sigma_P^2 = 49/3$ (Fig. 4.2C): C-fixation and D-fixation are both SLS which shows that C and D can be simultaneously favored by natural selection.

Example 4. $\sigma_R^2 = \sigma_S^2 = 0$, $\sigma_T^2 = 100/3$ and $\sigma_P^2 = 64/3$ (Fig. 4.2D): C-fixation is SLS but D-fixation is SLU which shows that only C can be favored by natural selection.

Obviously, all these simulation results not only match the theoretical predictions but also show that stochasticity in the environment could be conducive to the evolution and maintenance of cooperation.

Now, let us go back to the RPD game introduced in the beginning of this paper,

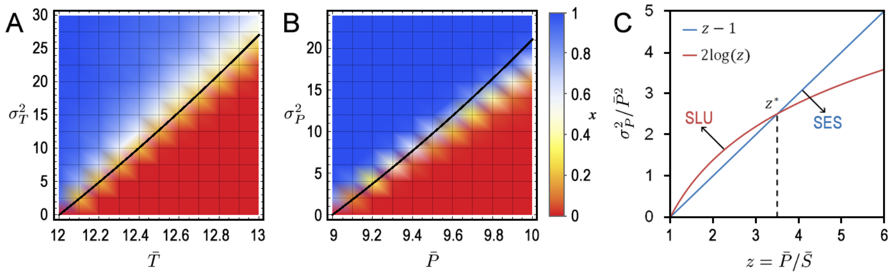


Figure 4.1: *The stochastic local stability of C-fixation and D-fixation and the stochastic evolutionary stability of strategy D.* (**A**) The black curve represents the function $\sigma_T^2 = -2\bar{T}^2 \log\left(\bar{R}/\bar{T}\right)$, where $\bar{R} = 12$, $\bar{P} = 10$, $\bar{S} = 9$, $\sigma_R^2 = \sigma_P^2 = \sigma_S^2 = 0$, and $T_t = \bar{T} + \delta_t$ with δ_t being a random variable uniformly distributed on $[-\delta, \delta]$ so that $\sigma_T^2 = \delta^2/3$. The C-fixation is SLS (or SLU) if σ_T^2 is above (or under) the curve. The color of each point on the $\bar{T}$-σ_T^2 plane represents the average value of the frequency of C after 5×10^3 time steps over 10^3 runs starting with $x_0 = 0.99$. (**B**) The black curve represents the function $\sigma_P^2 = 2\bar{P}^2 \log\left(\bar{P}/\bar{S}\right)$, where $\bar{R} = 12$, $\bar{S} = 9$, $\bar{T} = 13$, $\sigma_R^2 = \sigma_S^2 = 0$, $T_t = 13 + \delta'_t$ with δ'_t being a random variable uniformly distributed on $[-10, 10]$ so that $\sigma_T^2 = 100/3$, and $P_t = \bar{P} + \delta''^2_t$ with δ''^2_t being a random variable uniformly distributed on $[-\delta'', \delta'']$ so that $\sigma_P^2 = \delta''^2/3$. The D-fixation is SLU (or SLS) if σ_P^2 is above (or under) the curve. The color of each point on the $\bar{P}$-σ_P^2 plane also represents the average value of the frequency of C after 5×10^3 time steps over 10^3 runs starting with $x_0 = 0.01$. (**C**) The blue line represents the function $\sigma_P^2/\bar{P}^2 = z - 1$ with $z = \bar{P}/\bar{S}$, and the red curve the function $\sigma_P^2/\bar{P}^2 = 2\log(z)$. The blue line and the red curve intersects at $z = z^*$ so that (i) D-fixation is neither SLS nor SES (or is both SLS and SES) if $\sigma_P^2/\bar{P}^2$ is above (or under) both the blue line and the red curve; and (ii) D-fixation is SES (or SLS) but not SLS (nor SES) if $\sigma_P^2/\bar{P}^2$ is under the blue line (or the red curve) but above the red curve (or the blue line).

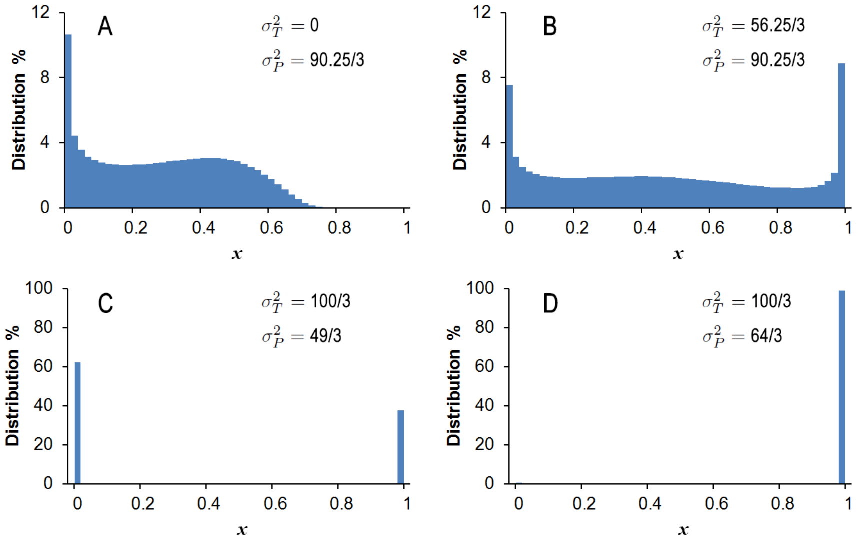


Figure 4.2: *The probability distribution of the frequency of cooperation in simulations.* Panels (**A**), (**B**), (**C**) and (**D**) correspond to the simulation results in examples 1, 2, 3 and 4 in the text, respectively, with the same mean random payoffs $\bar{R} = 12$, $\bar{S} = 9$, $\bar{T} = 13$ and $\bar{P} = 10$.

which is an extension of the classic Tucker's PD game (Poundstone, 1992; Nowak and Highfield, 2011). Consider the random payoff matrix at time step $t \geq 0$ with entries $R_t = u + a_{11}(t)$, $S_t = u + a_{12}(t)$, $T_t = u + a_{21}(t)$ and $P_t = u + a_{22}(t)$, where $a_{11}(t) = -1, -2, -3$ with probability distribution $\mathbf{P}_{11} = (1/3, 1/3, 1/3)$, $a_{12}(t) = -3, -4, -5$ with $\mathbf{P}_{12} = \mathbf{P}_{11}$, $a_{21}(t) = 0, -1, -2$ with $\mathbf{P}_{21} = \mathbf{P}_{11}$, and $a_{22}(t) = -2, -3, -4$ with $\mathbf{P}_{22} = \mathbf{P}_{11}$. Here, $u \geq 5$ is a positive constant such that $u + a_{ij}(t) \geq 0$ for $i, j = 1, 2$. Then, C-fixation is SLU, while D-fixation is SLS and strategy D is SES. However, if we take $a_{22}(t) = 0, -3, -6$ with probability distribution $\mathbf{P}_{22} = (1/2, 0, 1/2)$ and $u = 6$, then D-fixation is SLU and strategy D is not SES.

4.3 Conclusion

In the classic PD game with a constant payoff matrix, if all members of a population adopt the defection strategy, then no other mutant strategy can successfully invade the population under the effects of natural selection (Axelrod, 1984; Hofbauer an Sigmund, 1998; Nowak, 2006). Stochastic fluctuations in the environment are generally considered to be among the main driven forces of evolution in biological populations and human societies (Kaplan *et al.*, 1990; Zheng *et al.*, 2017). However, it is still unclear how they could favor the appearance, maintenance and spread of cooperation in natural populations. Our objective in this study was to show how cooperation, at least when rate or widespread, could evolve or be maintained under the effects of random environmental noise. Extending the classic PD game with a constant payoff matrix to a randomized PD (RPD) game with a random payoff matrix having the same mean, the focus has been put on stochastic local stability (SLS) of the fixation states and stochastic evolutionary stability (SES) of the corresponding pure strategies, C for cooperation and D for defection. Our theoretical results supported by simulations suggest that a decrease of uncertainty in the payoffs to C (namely, a decrease in $\sigma_R/\bar{R}$ and $\sigma_S/\bar{S}$ for the payoffs to C against C and D, respectively), or an increase of uncertainty in the payoffs to D (namely, $\sigma_T/\bar{T}$ and $\sigma_P/\bar{P}$ for the payoffs to D against C and D, respectively), could be propitious to the evolution of cooperation. All these results not only reveal how the evolutionary dynamics in the Prisoner's Dilemma game may be influenced by stochasticity but also provide a new perspective for a better understanding of the evolution of cooperative behavior.

4.4 Appendix

4.4.1 A. Stochastic local stability

Consider a two-phenotype game in an infinite population with discrete, non-overlapping generations. For convenience, two strategies are denoted by 1 and 2, respectively. The payoffs in pairwise interactions at time step $t \geq 0$ are given by the payoff matrix

$$\mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix}, \tag{4.4.1}$$

where $a_{ij}(t)$ is the payoff to strategy i against strategy j at time step $t \geq 0$ for $i, j = 1, 2$. Without loss of generality, these payoffs are assumed to be positive random variables, and we further assume that there exist real numbers $A, B > 0$ with $A < B$ such that $A \leq a_{ij}(t) \leq B$ for $i, j = 1, 2$ and all $t \geq 0$. Moreover, the probability distributions of $a_{ij}(t)$ for $i, j = 1, 2$ are the same for all $t \geq 0$, and the means, variances and covariances of these random payoffs are denoted by $\langle a_{ij}(t) \rangle = \bar{a}_{ij}$, $\left\langle \big(a_{ij}(t) - \bar{a}_{ij}\big)^2 \right\rangle = \sigma_{ij}^2$ and $\big\langle \big(a_{ij}(t) - \bar{a}_{ij}\big)\big(a_{kl}(t) - \bar{a}_{kl}\big) \big\rangle = \sigma_{ij,kl}$, respectively, for $i, j, k, l = 1, 2$. The payoffs $a_{ij}(s)$ and $a_{kl}(t)$ for $i, j, k, l = 1, 2$ are assumed to be independent of each other for $s, t \geq 0$ with $s \neq t$, in which case $\big\langle \big(a_{ij}(s) - \bar{a}_{ij}\big)\big(a_{kl}(t) - \bar{a}_{kl}\big) \big\rangle = 0$.

Let x_t be the frequency of strategy 1 at time step $t \geq 0$, and $1 - x_t$ the frequency of strategy 2. The expected payoff of strategy i at time step $t \geq 0$ is given by $\pi_{i,t} = x_t a_{i1}(t) + (1 - x_t) a_{i2}(t)$ for $i = 1, 2$, and the corresponding average payoff of the population is $\bar{\pi}_t = x_t \pi_{1,t} + (1 - x_t) \pi_{2,t}$. Moreover, the frequency of strategy 1 from time step t to time step $t+1$ is given by the recurrence equation

$$\begin{aligned} x_{t+1} &= \frac{x_t \pi_{1,t}}{\bar{\pi}_t} \\ &= \frac{x_t\big(x_t a_{11}(t) + (1-x_t)a_{12}(t)\big)}{x_t\big(x_t a_{11}(t) + (1-x_t)a_{12}(t)\big) + (1-x_t)\big(x_t a_{21}(t) + (1-x_t)a_{22}(t)\big)} \end{aligned} \tag{4.4.2}$$

for all $t \geq 0$ (Zheng *et al.*, 2017). Define $u_t = x_t\big/(1-x_t)$, the recurrence equation takes the form

$$u_{t+1} = u_t \left[\frac{u_t a_{11}(t) + a_{12}(t)}{u_t a_{21}(t) + a_{22}(t)}\right] \tag{4.4.3}$$

for all $t \geq 0$ (Zheng *et al.*, 2017).

Let $\hat{x}$ represent a constant equilibrium of Eq. (4.4.2), that is, an equilibrium of Eq. (4.4.2) that does not depend on the randomness of the payoff matrix $\mathbf{A}(t)$. Following Karlin and Liberman (1974, 1975), a constant equilibrium $\hat{x}$ is said to be stochastically locally stable (SLS) if for any $\epsilon > 0$ there exists $\delta_0 > 0$ such that $\mathbb{P}\big(x_t \to \hat{x}\big) \geq 1-\epsilon$ as soon as $|x_0 - \hat{x}| < \delta_0$. On the other hand, a constant equilibrium $\hat{x}$ is said to be stochastically locally unstable (SLU) if $\mathbb{P}\big(x_t \to \hat{x}\big) = 0$ as soon as $|x_0 - \hat{x}| > 0$. Here, we consider only the stochastic local stability of the boundary equilibria $\hat{x} = 0$ and $\hat{x} = 1$, and we leave aside the finer details of the mathematical analysis (the more rigorous proofs are similar to those in Zheng *et al.*, 2017).

Consider first the stochastic local stability of the boundary equilibrium $\hat{x} = 0$, which corresponds to the equilibrium $\hat{u} = \hat{x}\big/(1-\hat{x})$ in Eq. (4.4.3). Notice that this recurrence equation can be rewritten in the form

$$\frac{u_{t+1}}{u_t} = \frac{a_{12}(t)}{a_{22}(t)}\left[1 + \frac{u_t\big(a_{11}(t)a_{22}(t) - a_{12}(t)a_{21}(t)\big)}{u_t a_{12}(t)a_{21}(t) + a_{12}(t)a_{22}(t)}\right]. \tag{4.4.4}$$

Then, we have

$$\begin{aligned} \frac{1}{n}\Big[\log u_n - \log u_0\Big] &= \frac{1}{n}\sum_{t=0}^{n-1} \log\left[\frac{a_{12}(t)}{a_{22}(t)}\right] \\ &+ \frac{1}{n}\sum_{t=0}^{n-1}\log\left[1 + \frac{u_t\big(a_{11}(t)a_{22}(t) - a_{12}(t)a_{21}(t)\big)}{u_t a_{12}(t)a_{21}(t) + a_{12}(t)a_{22}(t)}\right]. \end{aligned} \tag{4.4.5}$$

Notice also that the second term on the right-hand side of the above equation tends to 0 as $u_t \to 0$. Thus, the strong law of large numbers guarantees that the boundary equilibrium $\hat{x} = 0$ is SLS if

$$\left\langle \log\left[\frac{a_{22}(t)}{a_{12}(t)}\right]\right\rangle \;=\; \left\langle \log a_{22}(t)\right\rangle - \left\langle \log a_{12}(t)\right\rangle \;>\; 0\ , \tag{4.4.6}$$

and SLU if the inequality is reversed (Zheng *et al.*, 2017). If both σ_{12}^2 and σ_{22}^2 are not too large so that the approximations $\left\langle \log a_{12}(t)\right\rangle \approx \log\bar{a}_{12} - \sigma_{12}^2/2\bar{a}_{12}^2$ and $\left\langle \log a_{22}(t)\right\rangle \approx \log\bar{a}_{22} - \sigma_{22}^2/2\bar{a}_{22}^2$ hold, then $\hat{x} = 0$ is SLS if

$$\log\left[\frac{\bar{a}_{22}}{\bar{a}_{12}}\right] \;>\; \frac{1}{2}\left[\frac{\sigma_{22}^2}{\bar{a}_{22}^2} - \frac{\sigma_{12}^2}{\bar{a}_{12}^2}\right]\ , \tag{4.4.7}$$

and SLU if the inequality is reversed.

By symmetry, the boundary equilibrium $\hat{x} = 1$ is SLS if

$$\left\langle \log\left[\frac{a_{11}(t)}{a_{21}(t)}\right]\right\rangle \;>\; 0\ , \tag{4.4.8}$$

and SLU if the inequality is reversed (Zheng *et al.*, 2017). As previously, if both σ_{11}^2 and σ_{21}^2 are not too large, then $\hat{x} = 1$ is SLS if

$$\log\left[\frac{\bar{a}_{11}}{\bar{a}_{21}}\right] \;>\; \frac{1}{2}\left[\frac{\sigma_{11}^2}{\bar{a}_{11}^2} - \frac{\sigma_{21}^2}{\bar{a}_{21}^2}\right]\ , \tag{4.4.9}$$

and SLU if the inequality is reversed.

4.4.2 B. Stochastic evolutionary stability

Let $\mathbf{x} = (x, 1 - x)$ denote a mixed strategy so that an individual using $\mathbf{x}$ will display strategy 1 with probability x and strategy 2 with complementary probability $1 - x$. In this context, a strategy $\hat{\mathbf{x}} = (\hat{x}, 1 - \hat{x})$ is said to be stochastically evolutionarily stable (SES) if its fixation is SLS against any other strategy $\mathbf{x} \neq \hat{\mathbf{x}}$ at least nearby enough (Zheng *et al.*, 2017).

Notice that the payoff matrix corresponding to the interaction between $\mathbf{x}$ and $\hat{\mathbf{x}}$ is given by

$$\begin{aligned} \mathbf{B}(t) &= \begin{pmatrix} b_{11}(t) & b_{12}(t) \\ b_{21}(t) & b_{22}(t) \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{x}\mathbf{A}(t)\mathbf{x} & \mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}} \\ \hat{\mathbf{x}}\mathbf{A}(t)\mathbf{x} & \hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}} \end{pmatrix} \ , \end{aligned} \tag{4.4.10}$$

with

$$\begin{aligned} \bar{b}_{12} &= \big\langle \mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}} \big\rangle \\ &= x\big(\hat{x}\bar{a}_{11} + (1-\hat{x})\bar{a}_{12}\big) + (1-x)\big(\hat{x}\bar{a}_{21} + (1-\hat{x})\bar{a}_{22}\big) \ , \\ \bar{b}_{22} &= \big\langle \hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}} \big\rangle \\ &= \hat{x}\big(\hat{x}\bar{a}_{11} + (1-\hat{x})\bar{a}_{12}\big) + (1-\hat{x})\big(\hat{x}\bar{a}_{21} + (1-\hat{x})\bar{a}_{22}\big) \ , \end{aligned}$$

and

$$\begin{aligned} \sigma^2_{b_{12}} &= \Big\langle \big[\mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}} - \big\langle \mathbf{x}\mathbf{A}(t)\hat{\mathbf{x}} \big\rangle\big]^2 \Big\rangle \\ &= x^2\hat{x}^2\sigma^2_{11} + x^2(1-\hat{x})^2\sigma^2_{12} + (1-x)^2\hat{x}^2\sigma^2_{21} + (1-x)^2(1-\hat{x})^2\sigma^2_{22} \\ &\quad +2\Big[x^2\hat{x}(1-\hat{x})\sigma_{11,12} + x(1-x)\hat{x}^2\sigma_{11,21} \\ &\quad +x(1-x)\hat{x}(1-\hat{x})(\sigma_{11,22} + \sigma_{12,21}) + x(1-x)(1-\hat{x})^2\sigma_{12,22} \\ &\quad +(1-x)^2\hat{x}(1-\hat{x})\sigma_{21,22}\Big] \ , \\ \sigma^2_{b_{22}} &= \Big\langle \big[\hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}} - \big\langle \hat{\mathbf{x}}\mathbf{A}(t)\hat{\mathbf{x}} \big\rangle\big]^2 \Big\rangle \\ &= \hat{x}^4\sigma^2_{11} + (1-\hat{x})^4\sigma^2_{22} + \hat{x}^2(1-\hat{x})^2\big(\sigma^2_{12} + \sigma^2_{21}\big) \\ &\quad +2\hat{x}(1-\hat{x})\Big[\hat{x}^2\big(\sigma_{11,12} + \sigma_{11,21}\big) + (1-\hat{x})^2\big(\sigma_{12,22} + \sigma_{21,22}\big) \\ &\quad +\hat{x}(1-\hat{x})\big(\sigma_{11,22} + \sigma_{12,21}\big)\Big] \ . \end{aligned}$$

Thus, for $\hat{\mathbf{x}} = (0,1)$, we have

$$\begin{aligned} \bar{b}_{12} &= x\bar{a}_{12} + (1-x)\bar{a}_{22} \ , \\ \bar{b}_{22} &= \bar{a}_{22} \ , \\ \sigma^2_{b_{12}} &= x^2\sigma^2_{12} + (1-x)^2\sigma^2_{22} + 2x(1-x)\sigma_{12,22} \ , \\ \sigma^2_{b_{22}} &= \sigma^2_{22} \ . \end{aligned} \tag{4.4.11}$$

Therefore, with small enough variances of the payoffs, the fixation state $\hat{\mathbf{x}} = (0,1)$ is SLS if

$$\log\left[\frac{x\bar{a}_{12} + (1-x)\bar{a}_{22}}{\bar{a}_{22}}\right] < \frac{1}{2}\left[\frac{x^2\sigma^2_{12} + (1-x)^2\sigma^2_{22} + 2x(1-x)\sigma_{12,22}}{(x\bar{a}_{12} + (1-x)\bar{a}_{22})^2} - \frac{\sigma^2_{22}}{\bar{a}^2_{22}}\right] . \tag{4.4.12}$$

Letting $x \to 0$ (but $x \neq 0$), we conclude that $\hat{\mathbf{x}} = (0,1)$ is SES if

$$\begin{aligned} \frac{\bar{a}_{12} - \bar{a}_{22}}{\bar{a}_{22}} &< -\frac{\big(\sigma^2_{22} - \sigma_{12,22}\big)\bar{a}_{22} + \big(\bar{a}_{12} - \bar{a}_{22}\big)\sigma^2_{22}}{\bar{a}^3_{22}} \\ &= -\frac{\bar{a}_{12}\sigma^2_{22} - \bar{a}_{22}\sigma_{12,22}}{\bar{a}^3_{22}} \ . \end{aligned} \tag{4.4.13}$$

Similarly, for $\hat{\mathbf{x}} = (1,0)$, we have

$$\begin{aligned} \bar{b}_{12} &= x\bar{a}_{11} + (1-x)\bar{a}_{21} \ , \\ \bar{b}_{22} &= \bar{a}_{11} \ , \\ \sigma^2_{b_{12}} &= x^2\sigma^2_{11} + (1-x)^2\sigma^2_{21} + 2x(1-x)\sigma_{11,21} \ , \\ \sigma^2_{b_{22}} &= \sigma^2_{11} \ , \end{aligned} \tag{4.4.14}$$

and the fixation of $\hat{\mathbf{x}} = (1,0)$ is SLS if

$$\log\left[\frac{x\bar{a}_{11} + (1-x)\bar{a}_{21}}{\bar{a}_{11}}\right] < \frac{1}{2}\left[\frac{x^2\sigma^2_{11} + (1-x)^2\sigma^2_{21} + 2x(1-x)\sigma_{11,21}}{(x\bar{a}_{11} + (1-x)\bar{a}_{21})^2} - \frac{\sigma^2_{11}}{\bar{a}^2_{11}}\right] . \tag{4.4.15}$$

We conclude that $\hat{\mathbf{x}} = (1,0)$ is SES if

$$\begin{aligned} \frac{\bar{a}_{11} - \bar{a}_{21}}{\bar{a}_{11}} &> -\frac{\big(\sigma^2_{11} - \sigma_{11,21}\big)\bar{a}_{11} + \big(\bar{a}_{11} - \bar{a}_{21}\big)\sigma^2_{11}}{\bar{a}^3_{11}} \\ &= -\frac{\bar{a}_{21}\sigma^2_{11} - \bar{a}_{11}\sigma_{11,21}}{\bar{a}^3_{11}} \ . \end{aligned} \tag{4.4.16}$$

Acknowledgements

Funding: In this study, T.J., Q.H., X.Z. and Y.T. were supported by the National Natural Science Foundation of China (Grants No.31971511 and No.31770426) and the Ministry of Science and Technology of China (Grant No.2018YFC1003300); S.L. was supported by Chinese Academy of Sciences President's International Fellowship Initiative (Grant No. 2016VBA039); C.L. and S.L. were supported in part by Natural Sciences and Engineering Research Council of Canada, Grant No.8833.

Author contributions: C.L., X.Z., S.L. and Y.T. performed theoretical analysis. C.L. ran the simulations. Y.T. wrote the early draft. C.L, T.J., Q.H., X.Z., S.L., and Y.T. reviewed the manuscript. C.L., T.J., and Q.H. contributed equally to this work.

Competing interests: Authors declare no competing interests.

Bibliography

[1] W. Poundstone, *Prisoner's Dilemma* (Doubleday, New York, 1992).

[2] M.A. Nowak, R. Highfield, *Supercooperators: altruism, evolution, and why we need each other to succeed* (Free Press, New York, 2011).

[3] J. Hofbauer, K. Sigmund, *The Theory of Evolution and Dynamical Systems* (Cambridge University Press, Cambridge, 1998).

[4] M.A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).

[5] D. Fudenberg, D.G. Rand, A. Dreber, Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review* **102**(2), 720-749 (2012).

[6] E. Solan, N. Vieille, Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America* **112**(45), 13743-13746 (2015).

[7] C. Hilbe, Š. Šimsa, K. Chatterjee, M.A. Nowak, Evolution of cooperation in stochastic games. *Nature* **559**, 246-249 (2018).

[8] R. Axelrod, W.D. Hamilton, The evolution of cooperation. *Science* **211**, 1390-1396 (1981).

[9] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).

[10] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[11] M.A. Nowak, K. Sigmund, How populations cohere: five rules for cooperation in *Theoretical Ecology - Principles and Applications*, chap. 2, R. May, A. McLean, Eds. (Oxford University Press, Oxford, 2007).

[12] R.L. Trivers, Evolution of reciprocal altruism. *Quarterly Review of Biology* **46**, 35-57 (1971).

[13] R. Lande, S. Engen, B.-E. Saether, *Stochastic Population Dynamics in Ecology and Conservation* (Oxford University Press, Oxford, 2003).

[14] R.M. May, *Stability and complexity in model ecosystems* (Princeton University Press, Princeton, 1973).

[15] X-D. Zheng, C. Li, S. Lessard, Y. Tao, Evolutionary stability concepts in a stochastic environment. *Physical Review E* **96**, 032414 (2017).

[16] X-D. Zheng, C. Li, S. Lessard, Y. Tao, Environmental noise could promote stochastic local stability of behavioral diversity evolution. *Physical Review Letters* **120**, 218101 (2018).

[17] S. Karlin, U. Liberman, Random temporal variation in selection intensities: case of large population size. *Theoretical Population Biology* **6**, 355-382 (1974).

[18] S. Karlin, U. Liberman, Random temporal variation in selection intensities: one-locus two-allele model. *Journal of Mathematical Biology* **2**, 1-17 (1975).

[19] W.J. Ewens, *Mathematical Population Genetics: I Theoretical Introduction* (Springer, New York, 2004).

[20] J. Maynard Smith, G.R. Price, The logic of animal conflict. *Nature* **246**, 15-18 (1973).

[21] J. Maynard Smith, The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology* **47**, 209-221 (1974).

[22] D. DP Johnson, P. Stopka, J. Bell, Individual variation evades the Prisoner's Dilemma. *BMC Evolutionary Biology* **2**, 15 (2002).

[23] M. Perc, Coherence resonance in a spatial prisoner's dilemma game. *New Journal of Physics* **8**, 22 (2006).

[24] A. Szolnoki, M. Perc, Seasonal payoff variations and the evolution of cooperation in social dilemmas. *Scientific Reports* **9**, 12575 (2019).

[25] P.D. Taylor, L. Jonker, Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* **40**, 145-156 (1978).

[26] S. Lessard, Evolutionary dynamics in frequency-dependent two-phenotype models. *Theoretical Population Biology* **25**, 210-234 (1984).

[27] H. Kaplan, K. Hill, A.M. Hurtado, Risk, foraging and food sharing among the Ache in *Risk and Uncertainty in Tribal and Peasant Economies*, E. A. Cashdan, Ed. (Westview Press, Boulder, CO, 1990).

Chapter 5

Article 5

Randomized matrix games in a finite population: effect of stochastic fluctuations in the payoffs on the evolution of cooperation

Cong Li and Sabin Lessard

Département de mathématiques et de statistique
Université de Montréal, Montréal, Québec
H3C 3J7 CANADA

Information

This paper appeared in *Theoretical Population Biology* **134**, 77-91. It was received on 4 December 2019; available online on 22 May 2020.

My contribution

Sabin Lessard initiated the project and provided some early calculations. I completed the theoretical framework and wrote the first draft of the paper.

Abstract

A diffusion approximation for a randomized 2×2-matrix game in a large finite population is ascertained in the case of random payoffs whose expected values, variances and covariances are of order given by the inverse of the population size N. Applying the approximation to a Randomized Prisoner's Dilemma (RPD) with independent payoffs for cooperation and defection in random pairwise interactions, conditions on the variances of the payoffs for selection to favor the evolution of cooperation, favor more the evolution of cooperation than the evolution of defection, and disfavor the evolution of defection are deduced. All these are obtained from probabilities of ultimate fixation of a single mutant. It is shown that the conditions are lessened with an increase in the variances of the payoffs for defection against cooperation and defection and a decrease in the variances of the payoffs for cooperation against cooperation and defection. A RPD game with independent payoffs whose expected values are additive is studied in detail to support the conclusions. Randomized matrix games with non-independent payoffs, namely the RPD game with additive payoffs for cooperation and defection based on random cost and benefit for cooperation and the repeated RPD game with Tit-for-Tat and Always-Defect as strategies in pairwise interactions with a random number of rounds, are studied under the assumption that the population-scaled expected values, variances and covariances of the payoffs are all of the same small enough order. In the first model, the conditions in favor of the evolution of cooperation hold only if the covariance between the cost and the benefit is large enough, while the analysis of the second model extends the results on the effects of the variances of the payoffs for cooperation and defection found for the one-round RPD game.

Running head: Randomized matrix games

Keywords and phrases: Cooperation. Diffusion approximation. Fixation probability. Prisoner's Dilemma. Random payoffs.

Mathematics Subject Classification (2010): Primary 92D25; Secondary 60J70

5.1 Introduction

Cooperative behavior is a phenomenon that is widely observed in nature. However, natural selection tends to enhance selfish behavior through fierce competition. In order to explain the rationality of cooperation and its evolution in natural populations, a two-player game known as the Prisoner's Dilemma (PD) has been widely studied as one of the most important theoretical frameworks (Axelrod and Hamilton, 1981; Maynard Smith, 1982; Axelrod, 1984; Poundstone, 1992; Nowak and Highfield, 2011). In an additive version of the PD game, cooperation takes the form of a donor who pays a cost c for a recipient to get a benefit b. Defection costs nothing and does not disqualify from receiving a benefit. Therefore, the payoff for cooperation never exceeds the payoff for defection (Nowak, 2006; Nowak and Sigmund, 2007). This is the case in more general versions of the PD game. Moreover, assuming random pairwise interactions in an infinite population and average payoffs as relative growth rates, the replicator equation (Taylor and Jonker, 1978) predicts global convergence to fixation of defection (Hofbauer and Sigmund, 1998).

In a finite population of constant size N undergoing discrete, non overlapping generations according to a Wright-Fisher model and more general models with exchangeable reproduction schemes (Fisher, 1930; Wright, 1931; Cannings, 1974; Ewens, 2004), the fixation probability for a neutral mutant type represented only once initially is just the inverse of the population size, that is, N^{-1}. If this probability becomes larger than N^{-1} in the presence of selection, then the mutant type has been said to be favored by selection (Nowak *et al.*, 2004). Several mechanisms have been considered to explain how cooperation could be favored by natural selection assuming additive effects of average payoffs on fitness (Nowak and Sigmund, 2007). This is the case, for instance, for cooperation taking the form of the "tit-for-tat" strategy (Trivers, 1971; Axelrod and Hamilton, 1981; Axelrod, 1984) starting with cooperation in a repeated PD game between randomly chosen partners if the number of rounds exceeds some threshold value (Nowak *et al.*, 2004). This is also the case in group-structured or graph-structured populations for modeling some social or geographical networks with local interactions (Ohtsuki *et al.*, 2006). However, with a one-round PD game and constant payoffs in a well-mixed population, the fitness of cooperation never exceeds the fitness of defection, and, as a result, cooperation cannot be favored by selection.

In nature, there are changes not only in the composition of a population but also in the surrounding environment in which the population finds itself. These can affect the payoffs that individuals receive as a result of interactions with others. Randomness in evolutionary games can take several forms such as probabilistic encounter rules or mixed strategies depending or not on the replies of others (Taylor and Jonker, 1978; Eshel and Cavalli-Sforza, 1982; Hofbauer and Sigmund, 1998). Of particular interest are stochastic games which allow the environment to change in response to the players' choices (Shapley, 1953; Fudenberg *et al.*, 2012; Solan and Vieille, 2015; Hilbe *et al.*, 2018). But also not to be forgotten are variations in payoffs caused by disturbances in the natural environment. These can be periodic, e.g., being seasonal or alternating day and night. But they can also be totally random as if occurring by accident (May, 1973; Kaplan *et al.*, 1990; Lande *et al.*, 2003). In the case of deterministically time-dependent payoffs in 2 matrix games, for instance, Broom (2005) compares the time average of the population state and the interior Nash equilibrium of the average payoff matrix and shows that they can be arbitrarily far apart. With periodic payoffs, even stable periodic orbits can be found from arbitrary starting points (Uyttendaele *et al.*, 2012). On the other hand, Stollmeier and Nagler (2018) shows that under the effects of random environmental noise, an evolutionary game involving two strategies with a strategy having a higher expected payoff at any frequency than the other can reach a stationary distribution with both strategies co-existing.

In a matrix game, unless stochastic fluctuations in the environment are small enough to be ignored, it is more accurate to use random payoffs than constant payoffs. In particular, the introduction of random payoffs extends the classical PD game to a randomized PD game. In order to reveal how environmental noise can generally affect the evolutionary game dynamics in an infinite population, the concepts of stochastic evolutionary stability (SES) and stochastic convergence stability (SCS) have been investigated (Zheng *et al.*, 2017, 2018). Applying these concepts to a one-round randomized PD game in a well-mixed population, it can be shown that the evolution of cooperation tends to be more easily favored by natural selection if the coefficients of variation of the payoffs are smaller for cooperation than for defection (Li *et al.*, 2019).

On the other hand, in a population genetics framework for a large finite population, Karlin and Levikson (1974) have shown that, when the mean and variance of frequency-independent genotypic fitnesses are of the same order given by the inverse of the population size, the effect of the variance matters. Actually, the variance in selection, meaning fluctuating selection intensities, produces a "drift effect" away from the fixation states.

In order to study the effect of stochastic fluctuations in a context of an evolutionary game in a large finite population, we consider in this paper a matrix game with random payoffs for two players using one of two strategies. After ascertaining a diffusion approximation for this model, we focus on the Randomized Prisoner's Dilemma (RPD) with cooperation and defection as strategies, and we consider the probability of ultimate fixation of either strategy as a single mutant. Conditions that favor the evolution of cooperation are examined in detail in the case of independent payoffs such that the average effects of cooperation and defection are additive. A RPD game with random additive effects of cooperation and defection on the payoffs as well as a repeated RPD game are also studied.

5.2 The Model

We consider a randomized matrix game with two strategies in a finite population of fixed finite size N. The two possible pure strategies used by the individuals in the population are denoted by S_1 and S_2. At time $t \geq 0$ corresponding to some generation, the frequencies of S_1 and S_2 are given by $x(t)$ and $1 - x(t)$, respectively, while their payoffs in pairwise interactions are given by the entries of the 2×2 random game matrix

$$\begin{pmatrix} \eta_1(t) & \eta_2(t) \\ \eta_3(t) & \eta_4(t) \end{pmatrix}. \tag{5.2.1}$$

Here, $\eta_1(t)$ and $\eta_2(t)$ are the payoffs to strategy S_1 against strategies S_1 and S_2, respectively, while $\eta_3(t)$ and $\eta_4(t)$ are the corresponding payoffs to strategy S_2 against the same two strategies. We assume that the value of these payoffs are random variables with values that are always larger than -1 and probability distributions that do not depend on time $t \geq 0$.

In addition, we assume that these payoffs have expected values, variances and covariances of order given by the inverse of the population size, which will be taken later on as the time interval between two successive generations (see below). More precisely, they can be expressed in the form

$$\eta_i(t) = \mu_i N^{-1} + \xi_i(t), \tag{5.2.2}$$

where $\mathbf{E}(\xi_i(t)) = 0$, $\mathbf{Var}(\xi_i(t)) = \sigma_i^2 N^{-1}$ and $\mathbf{Cov}(\xi_i(t), \xi_j(t)) = \sigma_{ij} N^{-1}$, for $i, j = 1, \ldots, 4$ with $i \neq j$. Therefore, we have

$$\mathbf{E}(\eta_i(t)) = \mu_i N^{-1}, \tag{5.2.3a}$$

$$\mathbf{E}(\eta_i(t)^2) = \sigma_i^2 N^{-1} + o(N^{-1}), \tag{5.2.3b}$$

$$\mathbf{E}(\eta_i(t)\eta_j(t)) = \sigma_{ij} N^{-1} + o(N^{-1}), \tag{5.2.3c}$$

so that μ_i, σ_i^2 and σ_{ij} represent population-scaled parameters for the expected value, variance and covariance of the payoffs, respectively, for $i, j = 1, \ldots, 4$ with $i \neq j$. Moreover, it is assumed that

$$\mathbf{E}(\xi_1(t)^k \xi_2(t)^l \xi_3(t)^m \xi_4(t)^n) = o(N^{-1}), \tag{5.2.4}$$

so that

$$\mathbf{E}(\eta_1(t)^k \eta_2(t)^l \eta_3(t)^m \eta_4(t)^n) = o(N^{-1}), \tag{5.2.5}$$

for non-negative integers k, l, m, n such that $k + l + m + n \geq 3$.

We suppose that the payoffs have additive effects on fitness understood as relative reproductive success with a baseline value equal to 1. Assuming random pairwise interactions, the mean fitness of strategy S_1 at time $t \geq 0$ can be expressed as

$$\mathbf{\Pi}_1(t) = 1 + x(t)\eta_1(t) + (1 - x(t))\eta_2(t), \tag{5.2.6}$$

and the corresponding mean fitness of strategy S_2 as

$$\mathbf{\Pi}_2(t) = 1 + x(t)\eta_3(t) + (1 - x(t))\eta_4(t). \tag{5.2.7}$$

Note that these quantities are always positive since we assume $\eta_i(t) > -1$ for $i = 1, \ldots, 4$.

Now, we consider discrete non-overlapping generations as in the Wright-Fisher model and we measure time in number of N generations. Then, $\Delta t = N^{-1}$ represents the time interval from one generation to the next. Given that the frequency of strategy S_1 is $x(t)$ at time $t \geq 0$ corresponding to some generation, the frequency of S_1 in the next generation, $x(t+\Delta t)$, is distributed as a binomial random variable divided by N. Actually the conditional probability distribution is given by

$$x(t + \Delta t)|x(t) \sim \frac{1}{N}\mathrm{B}(N, x'(t)), \tag{5.2.8}$$

where $\mathrm{B}(N, x'(t))$ denotes a binomial distribution of parameters N and $x'(t)$ with

$$x'(t) = \frac{x(t)\mathbf{\Pi}_1(t)}{x(t)\mathbf{\Pi}_1(t) + (1 - x(t))\mathbf{\Pi}_2(t)} \tag{5.2.9}$$

being the probability for an offspring to have been produced by an individual using strategy S_1 at time $t \geq 0$. Note that $x'(t)$ is a random variable even if the value of $x(t)$ is known, since $\mathbf{\Pi}_1(t)$ and $\mathbf{\Pi}_2(t)$ depend on the random payoffs $\eta_i(t)$ for $i = 1, \ldots, 4$.

5.3 Diffusion Approximation

Let $\Delta x = x(t + \Delta t) - x(t)$ be the change in the frequency of individuals that use strategy S_1 from time t to time $t + \Delta t$. Given $x(t) = x$, the first, second and fourth moments of

Δx can be calculated as (see **Appendix A** for details)

$$\mathbf{E}(\Delta x | x(t) = x) = m(x)\Delta t + o(\Delta t), \tag{5.3.1}$$

$$\mathbf{E}((\Delta x)^2 | x(t) = x) = v(x)\Delta t + o(\Delta t) \tag{5.3.2}$$

and

$$\mathbf{E}((\Delta x)^4 | x(t) = x) = o(\Delta t), \tag{5.3.3}$$

where

$$\begin{aligned} m(x) = x(1-x)\Big(&\mu_2 - \mu_4 + x\big(\mu_1 - \mu_2 - \mu_3 + \mu_4\big) \\ &+ x^3(\sigma_{13} - \sigma_1^2) + x(1-x)^2(2\sigma_{34} - \sigma_{14} - \sigma_{23} + \sigma_{24} - \sigma_2^2) \\ &+ x^2(1-x)(-2\sigma_{12} + \sigma_{14} + \sigma_{23} - \sigma_{13} + \sigma_3^2) + (1-x)^3(\sigma_4^2 - \sigma_{24})\Big) \end{aligned} \tag{5.3.4}$$

and

$$\begin{aligned} v(x) = x(1-x)\Big(&1 + x^3(1-x)(\sigma_1^2 + \sigma_3^2 - 2\sigma_{13}) + x(1-x)^3(\sigma_2^2 + \sigma_4^2 - 2\sigma_{24}) \\ &+ 2x^2(1-x)^2(\sigma_{12} + \sigma_{34} - \sigma_{14} - \sigma_{23})\Big). \end{aligned} \tag{5.3.5}$$

The above conditions ascertain a diffusion approximation with drift function $m(x)$ and diffusion function $v(x)$ in the limit of a large population with the population size N as unit of time (Kimura, 1964; Ewens, 2004).

In the diffusion approximation, the probability density function of S_1-frequency evaluated at x at time $t \geq 0$ given a value p at time 0, denoted by $f(x, p, t)$, satisfies the forward Kolmogorov (Fokker-Planck) equation

$$\frac{\partial f(x,p,t)}{\partial t} = -\frac{\partial}{\partial x}\left[m(x)f(x,p,t)\right] + \frac{\partial^2}{\partial x^2}\left[\frac{v(x)f(x,p,t)}{2}\right], \tag{5.3.6}$$

as well as the backward Kolmogorov equation

$$\frac{\partial f(x,p,t)}{\partial t} = m(p)\frac{\partial f(x,p,t)}{\partial p} + \frac{v(p)}{2}\frac{\partial^2 f(x,p,t)}{\partial p^2}. \tag{5.3.7}$$

Since no mutation is considered in the model at hand, the two boundaries $x = 0$ and $x = 1$ are absorbing states.

Let $u(p, t)$ denote the probability that strategy S_1 is fixed by time $t \geq 0$ so that $x(t) = 1$ given an initial frequency $x(0) = p$. This probability satisfies the backward Kolmogorov equation, that is,

$$\frac{\partial u(p,t)}{\partial t} = m(p)\frac{\partial u(p,t)}{\partial p} + \frac{v(p)}{2}\frac{\partial^2 u(p,t)}{\partial p^2} \tag{5.3.8}$$

with the boundary conditions $u(0, t) = 0$ and $u(1, t) = 1$. By letting $t \to \infty$, the limit

$$u(p) = \lim_{t\to\infty} u(p,t) \tag{5.3.9}$$

represents the probability of ultimate fixation of strategy S_1 given an initial frequency $x(0) = p$. As $t \to \infty$, the left-hand side in (5.3.8) tends to 0 so that we have

$$0 = m(p)\frac{du(p)}{dp} + \frac{v(p)}{2}\frac{d^2 u(p)}{dp^2} \tag{5.3.10}$$

with the boundary conditions $u(0) = 0$ and $u(1) = 1$. The solution of this ordinary differential equation is known to be (Risken, 1992; Ewens, 2004)

$$u(p) = \frac{\int_0^p \psi(y)dy}{\int_0^1 \psi(y)dy}, \tag{5.3.11}$$

where

$$\psi(y) = \exp\left(-2\int_0^y \frac{m(x)}{v(x)}dx\right). \tag{5.3.12}$$

Note that the probability of ultimate fixation of strategy S_2 is given by

$$1 - u(p) = \frac{\int_p^1 \psi(y)dy}{\int_0^1 \psi(y)dy}, \tag{5.3.13}$$

since there is ultimate fixation of strategy S_1 or S_2 with probability 1.

5.4 Randomized Prisoner's Dilemma (RPD)

Consider a random game matrix (5.2.1) with independent payoffs whose expected values determine a classical Prisoner's Dilemma (PD). In this case, the population-scaled parameters in (5.2.3) verify $\sigma_{ij} = 0$ for $i, j = 1, \ldots, 4$ with $i \neq j$, since the payoffs are uncorrelated, and

$$\begin{pmatrix} \mu_1 & \mu_2 \\ \mu_3 & \mu_4 \end{pmatrix} = \begin{pmatrix} R & S \\ T & P \end{pmatrix} \tag{5.4.1}$$

with $T > R > P > S$ and $2R > T + S$. Then, we have a randomized Prisoner's Dilemma (RPD) with strategies S_1 and S_2 corresponding to cooperation (C) and defection (D), respectively.

Suppose that cooperation is introduced as a single mutant in an all defecting population so that the initial frequency of cooperation in the population of size N is $p = N^{-1}$. If the probability of ultimate fixation of cooperation, denoted by $F_C = u(N^{-1})$, exceeds the value N^{-1}, which is the fixation probability under neutrality, then we say that *the evolution of cooperation is favored by selection.* Analogously, $F_D = 1 - u(1 - N^{-1})$ represents the probability of ultimate fixation of a single defecting mutant in an all cooperating population, and we say that *the evolution of defection is disfavored (not favored) by selection* if F_D is less than N^{-1} (Nowak *et al.*, 2004). Moreover, if $F_C > F_D$, then the invasion of a single cooperating mutant in an all defecting population is more likely than the reverse situation. In such a case, we say that *the evolution of cooperation is more favored by selection than the evolution of defection.* Finally, if all three conditions are satisfied, which occurs when $F_C > N^{-1} > F_D$, then we say that *the evolution of cooperation is fully favored by selection.*

Assuming the population size N large enough and using the diffusion approximation for the fixation probability, namely (5.3.11), the condition for the evolution of cooperation to be favored by selection becomes

$$F_C = \frac{\psi(0)}{N \int_0^1 \psi(y)dy} > \frac{1}{N}. \tag{5.4.2}$$

Since $\psi(0) = 1$, this condition is equivalent to

$$\int_0^1 \psi(y)dy < 1. \tag{5.4.3}$$

Moreover, owing to (5.3.4), (5.3.5) and (5.3.12), we have

$$\psi(y) = \exp\left(-2\int_0^y g(x)dx\right) \tag{5.4.4}$$

for $0 \leq y \leq 1$, where

$$g(x) = \frac{m(x)}{v(x)} \tag{5.4.5}$$

with drift function

$$\begin{aligned} m(x) =& x(1-x)\Big(\mu_2 - \mu_4 + x\big(\mu_1 - \mu_2 - \mu_3 + \mu_4\big) \\ & - \big(x^3\sigma_1^2 + x(1-x)^2\sigma_2^2 - x^2(1-x)\sigma_3^2 - (1-x)^3\sigma_4^2\big)\Big) \end{aligned} \tag{5.4.6}$$

and diffusion function

$$v(x) = x(1-x)\Big(1 + x^3(1-x)(\sigma_1^2 + \sigma_3^2) + x(1-x)^3(\sigma_2^2 + \sigma_4^2).\Big). \tag{5.4.7}$$

Note that the function $g(x)$ actually depends on the population-scaled expected values and variances of the payoffs, μ_i and σ_i^2 for $i = 1, \ldots, 4$.

Similarly, the condition for the evolution of defection not to be favored by selection takes the form

$$F_D = \frac{\psi(1)}{N\int_0^1 \psi(y)dy} = \frac{\phi(1)}{N\int_0^1 \phi(y)dy} < \frac{1}{N}, \tag{5.4.8}$$

where

$$\phi(y) = \frac{\psi(y)}{\psi(1)} = \exp\left(2\int_y^1 g(x)dx\right). \tag{5.4.9}$$

Since $\phi(1) = 1$, this condition is equivalent to

$$\int_0^1 \phi(y)dy > 1, \text{ that is, } \int_0^1 \psi(y)dy > \psi(1). \tag{5.4.10}$$

Moreover, since $F_D = \psi(1)F_C$, the condition

$$\psi(1) = \phi(0)^{-1} < 1 \tag{5.4.11}$$

ensures that $F_C > F_D$, in which case selection favors more the evolution of cooperation than the evolution of defection.

Let $h(x) = g(x)$ when $\sigma_i^2 = 0$ for $i = 1, \dots, 4$. Then we have

$$h(x) = \mu_2 - \mu_4 + x(\mu_1 - \mu_2 - \mu_3 + \mu_4). \tag{5.4.12}$$

Since $\mu_4 > \mu_2$ and $\mu_3 > \mu_1$ in the PD game (5.4.1), the function $h(x)$ is always negative for $0 \leq x \leq 1$. Therefore, in the case where $\sigma_i^2 = 0$ for $i = 1, \dots, 4$, we have $g(x) = h(x) < 0$ in (5.4.4) so that $\psi(y)$ and $\phi(y)$ are both strictly increasing functions with respect to y with $\psi(0) = \phi(1) = 1$. Therefore, we have $\psi(y) > 1$ and $\phi(y) < 1$ for $0 < y < 1$. In this case, conditions (5.4.3), (5.4.10) and (5.4.11) can never be satisfied. This means that the evolution of cooperation can never be favored by selection. This is in agreement with what is known for the classical PD game with deterministic payoffs (Maynard Smith, 1982; Nowak, 2006).

For the RPD game with independent payoffs, we consider the partial derivatives of $g(x)$ with respect to the variances of the payoffs. It can be shown (see **Appendix B** for details) that

$$\frac{\partial g(x)}{\partial \sigma_3^2} > 0 \tag{5.4.13}$$

and

$$\frac{\partial g(x)}{\partial \sigma_4^2} > 0 \tag{5.4.14}$$

for $0 < x < 1$. This implies that $g(x)$ for $0 < x < 1$ increases as σ_3^2 or σ_4^2 increases. Therefore, $\psi(y)$ in (5.4.4) for $0 < y < 1$, and its integral from 0 to 1 in (5.4.2) and (5.4.3), decreases as σ_3^2 or σ_4^2 increases. On the other hand, $\phi(y)$ in (5.4.9) for $0 < y < 1$, and its integral from 0 to 1 in (5.4.8) and (5.4.10), increases as σ_3^2 or σ_4^2 increases.

Let us summarize our findings.

Conclusion 1: *In a RPD game with independent payoffs, increasing the variance of at least one payoff for defection, that is,* σ_3^2 *or* σ_4^2*, increases the probability of ultimate* *fixation of cooperation introduced as a single mutant in an all defecting population, F_C, while it decreases the probability of ultimate fixation of defection introduced as a single mutant in an all cooperating population, F_D.*

5.5 RPD with independent payoffs

In this section, we focus on a RPD game with independent payoffs whose expected values are such that

$$\begin{pmatrix} \mu_1 & \mu_2 \\ \mu_3 & \mu_4 \end{pmatrix} = \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} . \tag{5.5.1}$$

This payoff matrix determines an additive PD game in which cooperation (C) incurs a fixed cost $c > 0$ to the individual adopting it, but provides a fixed benefit $b > 0$ to the opponent, while defection (D) incurs no cost at all.

In this case, the function $h(x)$ in (5.4.12) is given by $h(x) = -c$. Moreover, if $c \leq 1$, then it can be shown (see **Appendix B** for details) that the function $g(x)$ in (5.4.5) satisfies

$$\frac{\partial g(x)}{\partial \sigma_1^2} < 0 \tag{5.5.2}$$

and

$$\frac{\partial g(x)}{\partial \sigma_2^2} < 0 \tag{5.5.3}$$

for $0 < x < 1$.

This leads to the following complementary result.

Conclusion 2: *In a RPD game with independent payoffs whose population-scaled expected values determine an additive PD game in the form (5.5.1) with cost of cooperation $c \leq 1$, diminishing the variance of at least one payoff for cooperation, that is, σ_1^2 or σ_2^2, increases the probability of ultimate fixation of cooperation introduced as a single mutant in an all defecting population, F_C, while it decreases the probability of ultimate fixation of defection introduced as a single mutant in an all cooperating population, F_D.*

In the rest of this section, we investigate some special cases of the RPD with additive expected payoffs to exhibit conditions under which the evolution of cooperation could be favored by selection.

5.5.1 Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 0$, $\sigma_4^2 = \sigma^2 > 0$

This is a situation where the variance of the payoff for defection against defection is significantly larger than the variances of all the other payoffs.

With $h(x) = -c$, the function $g(x)$ in (5.4.5) takes the form

$$g_1(x) := g(x) = \frac{-c + (1-x)^3\sigma^2}{1 + x(1-x)^3\sigma^2}. \tag{5.5.4}$$

This function satisfies $g(0) = \sigma^2 - c$, $g(1) = -c$, $g'(0) = -\sigma^2(3 + \sigma^2 - c)$ and $g'(1) = 0$ (see **Appendix C** for details).

If $\sigma^2 < c$, then $g(x) < 0$ for $0 < x < 1$. In this case, the function $\psi(y)$ in (5.4.4) satisfies $\psi(y) > 1$ for $0 < y \leq 1$, which entails $\int_0^1 \psi(y)dy > 1$, that is, $F_C < 1/N$.

On the other hand, if $\sigma^2 > c$, then we have $g'(x) < 0$ for $0 < x < 1$. In this case, $g(x)$ is a strictly decreasing function from $\sigma^2 - c > 0$ at $x = 0$ to $-c < 0$ at $x = 1$, while $\psi(y)$ is a strictly convex function for $0 \leq y \leq 1$. The unique point x^* between 0 and 1 where $g(x)$ crosses the x axis is the global minimum point of $\psi(y)$ for $0 \leq y \leq 1$ (see Figure 5.1a,b). Since $\psi(0) = 1$ with $\psi'(0) = -2g(0) = -2(\sigma^2 - c) < 0$, the condition $\psi(1) < 1$, which implies $F_C > F_D$, guarantees also that $\int_0^1 \psi(y)dy < 1$, which implies $F_C > 1/N$.

Let σ_{**}^2 be the value of $\sigma^2 > c$ such that $\int_0^1 g(x)dx = 0$, that is, $\psi(1) = 1$. Recall that $g(x)$ is strictly increasing as a function of σ^2 owing to (5.4.14). Consequently, the condition $\sigma^2 > \sigma_{**}^2$ is necessary and sufficient to have $\int_0^1 g(x)dx > 0$, that is, $\psi(1) < 1$, which implies $F_C > F_D$.

Now, let σ_*^2 be the value of σ^2 strictly comprised between c and σ_{**}^2 such that $\int_0^1 \psi(y)dy = 1$. Then we have $\int_0^1 \psi(y)dy < 1$ for $\sigma_*^2 < \sigma^2 < \sigma_{**}^2$. We conclude that $F_C > 1/N$ as soon as $\sigma^2 > \sigma_*^2 > c$.

Finally, let σ^2_{***} be the value of $\sigma^2 > \sigma^2_{**}$ such that $\int_0^1 \phi(y)dy = 1$. Then we have $\int_0^1 \phi(y)dy > 1$ for $\sigma^2 > \sigma^2_{***}$. This means that $F_D < 1/N$ if and only if $\sigma^2 > \sigma^2_{***}$ (see **Appendix E** for details).

If $\sigma^2/16$ is small, in which case $x(1-x)^3\sigma^2$ is small for $0 \leq x \leq 1$, we have the approximation

$$g(x) \approx -c + (1-x)^3\sigma^2 \tag{5.5.5}$$

for $0 \leq x \leq 1$. Moreover, using the approximation $e^z \approx 1 + z$ for z small enough, we get

$$\begin{aligned} \psi(y) &= \exp\left(-2\int_0^y g(x)dx\right) \\ &\approx 1 + 2cy - 2\sigma^2\left(y - \frac{3}{2}y^2 + y^3 - \frac{1}{4}y^4\right) \end{aligned} \tag{5.5.6}$$

and

$$\begin{aligned} \phi(y) &= \exp\left(2\int_y^1 g(x)dx\right) \\ &\approx 1 + 2c(y-1) - 2\sigma^2\left(y - \frac{3}{2}y^2 + y^3 - \frac{1}{4}y^4 - \frac{1}{4}\right) \end{aligned} \tag{5.5.7}$$

for $0 \leq y \leq 1$. Then, $\psi(1) = 1$ when

$$\sigma^2 = \sigma^2_{**} \approx 4c, \tag{5.5.8}$$

while $\int_0^1 \psi(y)dy = 1$ when

$$\sigma^2 = \sigma^2_{*} \approx \frac{5c}{2} \tag{5.5.9}$$

and $\int_0^1 \phi(y)dy = 1$ when

$$\sigma^2 = \sigma^2_{***} \approx 10c. \tag{5.5.10}$$

Here, we have $c < \sigma^2_* < \sigma^2_{**} < \sigma^2_{***}$ with $F_C > N^{-1}$, $F_C > F_D$ and $F_D < N^{-1}$ when $\sigma^2 > \sigma^2_*$, $\sigma^2 > \sigma^2_{**}$ and $\sigma^2 > \sigma^2_{***}$, respectively (see Figure 5.2 for a schematic representation of the situation and Table 5.1 for some particular values).

Our result suggests that the evolution of cooperation tends to be fully favored by selection with an increase of the variance of the payoff for defection against defection.

5.5.2 Case 2: $\sigma_1^2 = \sigma_2^2 = \sigma_4^2 = 0$, $\sigma_3^2 = \sigma^2 > 0$

Here, the variance of the payoff for defection against cooperation is significantly larger than the variances of all the other payoffs.

In this case, the function $g(x)$ in (5.4.5) becomes

$$g_2(x) := g(x) = \frac{-c + x^2(1-x)\sigma^2}{1 + x^3(1-x)\sigma^2}. \tag{5.5.11}$$

This function satisfies $g(0) = -c$, $g(1) = -c$, $g'(0) = 0$ and $g'(1) = -(c+1)\sigma^2$.

Note that $x^2(1-x) \leq 4/27$ for $0 \leq x \leq 1$, so that $g(x) \leq 0$ for $0 \leq x \leq 1$ if $\sigma^2 \leq (27/4)c$. Proceeding as in the previous case, this entails $\int_0^1 \psi(y)dy > 1$, that is, $F_C < 1/N$. Actually, this inequality is reversed only when $\sigma^2 > \sigma_*^2 > (27/4)c$, where σ_*^2 is the value of σ^2 such that $\int_0^1 \psi(y)dy = 1$.

If $\sigma^2/16$ is small, then

$$g(x) \approx -c + x^2(1-x)\sigma^2 \tag{5.5.12}$$

for $0 \leq x \leq 1$, from which

$$\psi(y) \quad \approx \quad 1 + 2cy - 2\sigma^2\left(\frac{1}{3}y^3 - \frac{1}{4}y^4\right) \tag{5.5.13}$$

and

$$\phi(y) \quad \approx \quad 1 + 2c(y-1) - 2\sigma^2\left(\frac{1}{3}y^3 - \frac{1}{4}y^4 - \frac{1}{12}\right) \tag{5.5.14}$$

for $0 \leq y \leq 1$. Then, $\psi(1) = 1$ when

$$\sigma^2 = \sigma_{**}^2 \approx 12c, \tag{5.5.15}$$

while $\int_0^1 \psi(y)dy = 1$ when

$$\sigma^2 = \sigma_*^2 \approx 15c \tag{5.5.16}$$

and $\int_0^1 \phi(y)dy = 1$ when

$$\sigma^2 = \sigma_{***}^2 \approx 10c. \tag{5.5.17}$$

Note that σ_*^2 and σ_{**}^2 are larger in case 2 than in case 1 and satisfy the inequalities $\sigma_*^2 > \sigma_{**}^2 > \sigma_{***}^2 > c$ (see Figure 5.1c,d and Table 5.1). The conditions for $F_C > 1/N$, $F_C > F_D$ and $F_D < 1/N$ remain the same as in the previous case, that is, $\sigma^2 > \sigma_*^2$, $\sigma^2 > \sigma_{**}^2$ and $\sigma^2 > \sigma_{***}^2$, respectively, but these conditions hold in a reverse order as the variance of the payoff for defection against cooperation increases.

We conclude that selection tends to fully favor the evolution of cooperation when the variance of the payoff for defection against cooperation increases.

5.5.3 Case 3: $\sigma_1^2 = \sigma_2^2 = 0$, $\sigma_3^2 = \sigma_4^2 = \sigma^2 > 0$

This is a situation where the payoffs for defection have a certain level of uncertainty while the payoffs for cooperation are much less variable.

In this case, the function $g(x)$ in (5.4.5) takes the form

$$g_3(x) := g(x) = \frac{-c + (1-x)\left(x^2 + (1-x)^2\right)\sigma^2}{1 + x(1-x)\left(x^2 + (1-x)^2\right)\sigma^2} \,. \tag{5.5.18}$$

This function satisfies $g(0) = \sigma^2 - c$, $g(1) = -c$, $g'(0) = -\sigma^2(3 + \sigma^2 - c)$ and $g'(1) = -(c+1)\sigma^2$.

Analogously to case 1, we have $g(x) < 0$ for $0 < x < 1$ when $\sigma^2 < c$. On the other hand, if $\sigma^2 > c$, then $g(x)$ is a decreasing function, while $\psi(y)$ and $\phi(y)$ are convex functions on $[0,1]$ (see **Appendix C** for details). Therefore, three threshold values of σ^2 could be found sequentially, that satisfy the inequality $c < \sigma_*^2 < \sigma_{**}^2 < \sigma_{***}^2$ (see **Appendix E** for details). As in case 1, if σ_*^2, σ_{**}^2 and σ_{***}^2 are the values of σ^2 such that $\int_0^1 \psi(y)dy = 1$, $\psi(1) = 1$ and $\int_0^1 \phi(y)dy = 1$, respectively, then $F_C > 1/N$, $F_C > F_D$ and $F_D < 1/N$ when $\sigma^2 > \sigma_*^2$, $\sigma^2 > \sigma_{**}^2$ and $\sigma^2 > \sigma_{***}^2$, respectively, with $c < \sigma_*^2 < \sigma_{**}^2 < \sigma_{***}^2$ (see **Appendix E** and Figures 5.1e,f and 5.2).

If $\sigma^2/16$ is small, we have the approximation

$$g(x) \approx -c + (1-x)\left(x^2 + (1-x)^2\right)\sigma^2 \tag{5.5.19}$$

for $0 \leq x \leq 1$, from which

$$\psi(y) \approx 1 + 2cy - 2\sigma^2 \left(y - \frac{3}{2}y^2 + \frac{4}{3}y^3 - \frac{1}{2}y^4 \right) \tag{5.5.20}$$

and

$$\phi(y) \approx 1 + 2c(y-1) - 2\sigma^2 \left(y - \frac{3}{2}y^2 + \frac{4}{3}y^3 - \frac{1}{2}y^4 - \frac{1}{3} \right) . \tag{5.5.21}$$

for $0 \leq y \leq 1$. Then, $\psi(1) = 1$ when

$$\sigma^2 = \sigma_{**}^2 \approx 3c, \tag{5.5.22}$$

while $\int_0^1 \psi(y)dy = 1$ when

$$\sigma^2 = \sigma_{*}^2 \approx \frac{15c}{7} \tag{5.5.23}$$

and $\int_0^1 \phi(y)dy = 1$ when

$$\sigma^2 = \sigma_{***}^2 \approx 5c. \tag{5.5.24}$$

Comparisons between numerical and approximate values are made in Table 5.1.

Note that

$$g_3(x) - g_1(x) = \frac{(1+cx)(1-x)x^2\sigma^2}{(1 + x(1-x)^3\sigma^2)\left(1 + x(1-x)\left(x^2 + (1-x)^2\right)\sigma^2\right)} > 0 \tag{5.5.25}$$

for $0 < x < 1$. Thus, we have

$$\psi_3(y) = \exp\left(-2\int_0^y g_3(x)dx\right) < \exp\left(-2\int_0^y g_1(x)dx\right) = \psi_1(y) \tag{5.5.26}$$

and

$$\phi_3(y) = \exp\left(2\int_y^1 g_3(x)dx\right) > \exp\left(2\int_y^1 g_1(x)dx\right) = \phi_1(y) \tag{5.5.27}$$

for $0 < y < 1$. This implies that the probability of ultimate fixation of cooperation (defection) introduced as a single mutant in case 3 is larger (smaller) than in case 1 for all values of $\sigma^2 > 0$. Moreover, the values σ_*^2, σ_{**}^2 and σ_{***}^2 are smaller in case 3 than in case 1.

In conclusion, increasing the variance in both σ_3^2 and σ_4^2 is always more favorable for the evolution of cooperation than increasing the variance in only one of them.

5.5.4 Case 4: $\sigma_1^2 = \sigma_2^2 = \sigma_0^2 > 0$, $\sigma_3^2 = \sigma_4^2 = \sigma^2 > 0$

Here, this is an example where the variances of the payoffs for cooperation are fixed while the variances of the payoffs for defection are changing.

With the given variances, the function $g(x)$ in (5.4.5) takes the form

$$g_4(x) := g(x) = \frac{-c + (1-x)\left(x^2 + (1-x)^2\right)\sigma^2 - x\left(x^2 + (1-x)^2\right)\sigma_0^2}{1 + x(1-x)\left(x^2 + (1-x)^2\right)\left(\sigma^2 + \sigma_0^2\right)}, \tag{5.5.28}$$

which satisfies $g(0) = \sigma^2 - c$, $g(1) = -c - \sigma_0^2$.

In this case, the functions $g(x)$, $\psi(y)$ (see Figure 5.1g,h) and $\phi(y)$, and the threshold values of σ^2, namely σ_*^2, σ_{**}^2 and σ_{***}^2, have the same properties as in cases 1 and 3 (see Appendices C and E, and Figure 5.2). Moreover, if $(\sigma^2 + \sigma_0^2)/16$ is small, then we have the approximation

$$g(x) \approx -c + (1-x)\left(x^2 + (1-x)^2\right)\sigma^2 - x\left(x^2 + (1-x)^2\right)\sigma_0^2 \tag{5.5.29}$$

for $0 \leq x \leq 1$, from which

$$\psi(y) \approx 1 + 2cy - 2\sigma^2\left(y - \frac{3}{2}y^2 + \frac{4}{3}y^3 - \frac{1}{2}y^4\right) + 2\sigma_0^2\left(\frac{1}{2}y^2 - \frac{2}{3}y^3 + \frac{1}{2}y^4\right) \tag{5.5.30}$$

and

$$\begin{aligned}\phi(y) \approx 1 + 2c(y-1) - 2\sigma^2\left(y - \frac{3}{2}y^2 + \frac{4}{3}y^3 - \frac{1}{2}y^4 - \frac{1}{3}\right)\\ + 2\sigma_0^2\left(\frac{1}{2}y^2 - \frac{2}{3}y^3 + \frac{1}{2}y^4 - \frac{1}{3}\right)\end{aligned} \; . \tag{5.5.31}$$

for $0 \leq y \leq 1$. Then, $\psi(1) = 1$ when

$$\sigma^2 = \sigma_{**}^2 \approx 3c + \sigma_0^2, \tag{5.5.32}$$

while $\int_0^1 \psi(y)dy = 1$ when

$$\sigma^2 = \sigma_*^2 \approx \frac{15c + 3\sigma_0^2}{7} \tag{5.5.33}$$

and $\int_0^1 \phi(y)dy = 1$ when

$$\sigma^2 = \sigma_{***}^2 \approx \frac{15c + 7\sigma_0^2}{3}. \tag{5.5.34}$$

Some values are given in Table 5.1. Since the threshold values σ_*^2, σ_{**}^2 and σ_{***}^2 for σ^2 increase with σ_0^2, these results reveal that larger is the value of σ_0^2, larger must be the value of σ^2 for selection to favor the evolution of cooperation in any sense. Moreover, note that $\sigma_{***}^2 > \sigma_{**}^2 > \sigma_0^2$.

The main conclusion is that a higher level of uncertainty in the payoffs for defection than in the payoffs for cooperation is required for the evolution of cooperation to be fully favored by selection. This is somehow in agreement with results that can be found for the RPD in an infinite population (Li *et al.*, 2019).

5.5.5 Case 5: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma^2 > 0$

This is a situation where all the variances of the payoffs are of the same magnitude.

With all variances equal to σ^2, the function $g(x)$ in (5.4.5) takes the form

$$g_5(x) := g(x) = \frac{-c + (1-2x)\left(x^2 + (1-x)^2\right)\sigma^2}{1 + 2x(1-x)\left(x^2 + (1-x)^2\right)\sigma^2}, \tag{5.5.35}$$

which satisfies $g(0) = \sigma^2 - c$, $g(1) = -c - \sigma^2$.

In this case, results (5.4.13) and (5.4.14) can no longer determine the monotonicity of $g(x)$ with respect to σ^2. Actually, it may be an increasing function for x near 0, and a decreasing function for x near 1. Nevertheless, it can be shown that $\psi(y)$ is decreasing with respect to σ^2 for $y \in (0, 1]$, which guarantees the existence of σ_*^2, while σ_{**}^2 and σ_{***}^2 do not exist (see Appendices D and E, and Figure 5.1i,j).

If $\sigma^2/16$ is small, then we have the approximation

$$g(x) \approx -c + (1-2x)\left(x^2 + (1-x)^2\right)\sigma^2 \tag{5.5.36}$$

for $0 \leq x \leq 1$, from which

$$\psi(y) \approx 1 + 2cy - 2\sigma^2\left(y - 2y^2 + 2y^3 - y^4\right) \tag{5.5.37}$$

for $0 \leq y \leq 1$. Then, $\int_0^1 \psi(y)dy = 1$ when

$$\sigma^2 = \sigma_*^2 \approx \frac{15c}{4}. \tag{5.5.38}$$

	c=1, σ_0^2=0.5			c=0.5, σ_0^2=0.5		
	σ_*^2	σ_{**}^2	σ_{***}^2	σ_*^2	σ_{**}^2	σ_{***}^2
Case 1	2.5 (2.66)	4 (4.34)	10 (10.92)	1.25 (1.29)	2 (2.08)	5 (5.21)
Case 2	15 (22.2)	12 (14.96)	10 (10.96)	7.5 (9.25)	6 (6.76)	5 (5.26)
Case 3	2.14 (2.13)	3 (2.95)	5 (4.91)	1.07 (1.07)	1.5 (1.49)	2.5 (2.47)
Case 4	2.36 (2.34)	3.5 (3.44)	6.17 (6.07)	1.29 (1.28)	2 (1.97)	3.66 (3.43)
Case 5	3.75 (3.67)			1.88 (1.86)		

Table 5.1: *Comparison between the numerical values and the approximate values of* σ_*^2*,* σ_{**}^2*,* σ_{***}^2 *in cases 1 to 5.* The population-scaled expected cost c for cooperation is set to 1 or 0.5. The value of σ_0^2 in case 4 is set to 0.5. The approximate values are given first followed by the numerical values in brackets.

Note that σ_*^2 is the only threshold value of σ^2 in this case (see Table 5.1 for particular values).

Therefore, the evolution of cooperation can be favored by selection, that is $F_C > N^{-1}$, but cannot be fully favored, which means that we cannot have $F_C > N^{-1} > F_D$. This is in agreement with the conclusion in case 4.

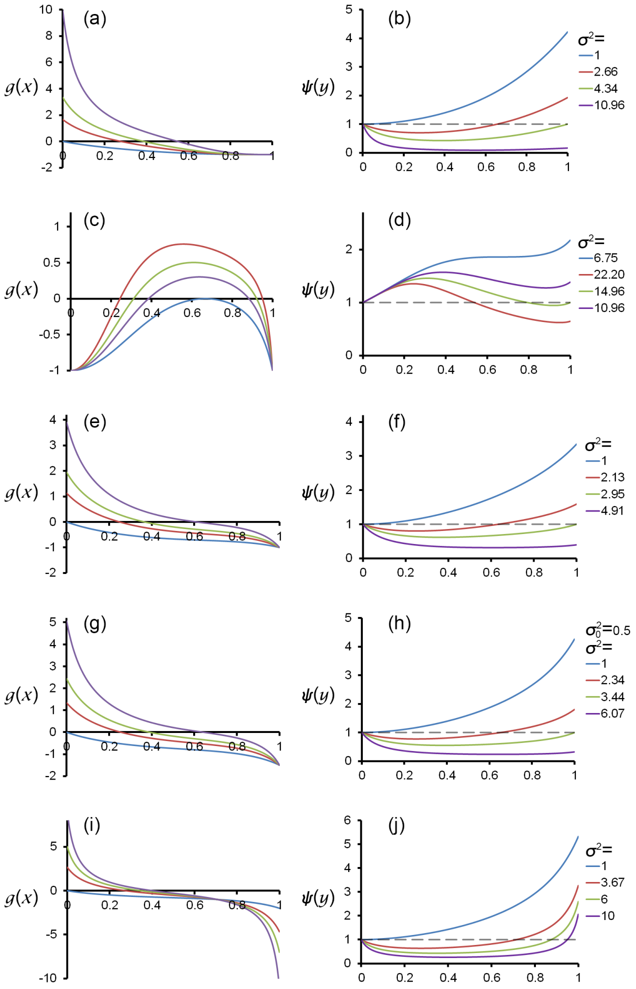


Figure 5.1: *Curves of* $g(x)$ *and* $\psi(y)$ *in cases 1 to 5.* The population-scaled expected cost c for cooperation is set to 1. In each panel, the curve in Blue is for the threshold value of σ^2 such that $g(x) \leq 0$ for $0 \leq x \leq 1$, while the curves in Red, Green and Purple are for the threshold values σ_*^2, σ_{**}^2 and σ_{***}^2, respectively (except for case 5 where there are no σ_{**}^2 and σ_{***}^2). Panels (a) and (b) represent $g(x)$ and $\psi(y)$ in case 1, and so on up to case 5. In case 4, the value of σ_0^2 is set to 0.5.

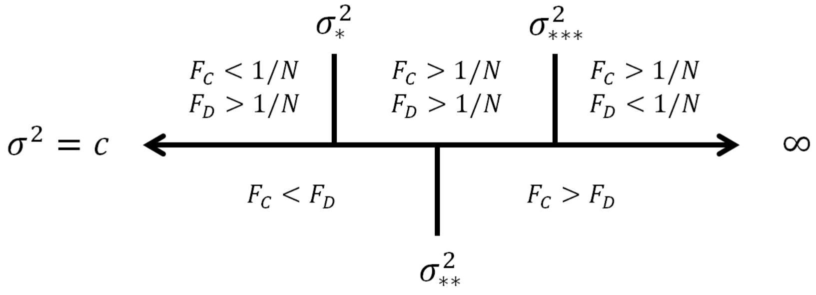


Figure 5.2: *Relationships between* σ^2_**,* σ^2_{**}*,* σ^2_{***} *and* F_C*,* F_D *in cases 1, 3 and 4.* The regions where the fixation probabilities F_C and F_D are larger or smaller than N^{-1} and where F_C is larger or smaller than F_D are given according to the position of σ^2 with respect to the increasing threshold values σ^2_*, σ^2_{**} and σ^2_{***}. In case 2, these threshold values are decreasing.

5.6 RPD with additive payoffs

In this section, we consider a RPD game with additive payoffs. At time $t \geq 0$, cooperation (C) incurs a random cost $c(t) > 0$ to the individual adopting it, but provides a random benefit $b(t) > 0$ to the opponent, while defection (D) incurs no cost at all, so that the random payoff matrix takes the form

$$\begin{pmatrix} \eta_1(t) & \eta_2(t) \\ \eta_3(t) & \eta_4(t) \end{pmatrix} = \begin{pmatrix} b(t) - c(t) & -c(t) \\ b(t) & 0 \end{pmatrix} . \tag{5.6.1}$$

The main difference with the model in the previous section is that the payoffs are not independent.

Here, $c(t)$ and $b(t)$ are assumed to be random variables with

$$\mathbf{E}(b(t)) = \mu_b N^{-1} > 0, \tag{5.6.2a}$$

$$\mathbf{E}(c(t)) = \mu_c N^{-1} > 0, \tag{5.6.2b}$$

$$\mathbf{Var}(b(t)) = \sigma_b^2 N^{-1}, \tag{5.6.2c}$$

$$\mathbf{Var}(c(t)) = \sigma_c^2 N^{-1}, \tag{5.6.2d}$$

$$\mathbf{Cov}(b, c) = \sigma_{bc} N^{-1}, \tag{5.6.2e}$$

so that the population-scaled parameters in (5.2.3) for the means, variances and covariances of the payoffs are given by

$$\mu_1 = \mu_b - \mu_c, \tag{5.6.3a}$$

$$\mu_2 = -\mu_c, \tag{5.6.3b}$$

$$\mu_3 = \mu_b, \tag{5.6.3c}$$

$$\mu_4 = 0, \tag{5.6.3d}$$

and

$$\sigma_1^2 = \sigma_b^2 + \sigma_c^2 - 2\sigma_{bc}, \tag{5.6.4a}$$

$$\sigma_2^2 = \sigma_c^2, \tag{5.6.4b}$$

$$\sigma_3^2 = \sigma_b^2, \tag{5.6.4c}$$

$$\sigma_{12} = \sigma_c^2 - \sigma_{bc}, \tag{5.6.4d}$$

$$\sigma_{13} = \sigma_b^2 - \sigma_{bc}, \tag{5.6.4e}$$

$$\sigma_{23} = -\sigma_{bc}, \tag{5.6.4f}$$

$$\sigma_{34} = \sigma_{24} = \sigma_{14} = \sigma_4^2 = 0. \tag{5.6.4g}$$

Substituting the above expressions into (5.3.4) and (5.3.5) yields

$$m(x) = x(1-x)\left(-\mu_c + x(\sigma_{bc} - \sigma_c^2)\right) \tag{5.6.5}$$

and

$$v(x) = x(1-x)\left(1 + x(1-x)\sigma_c^2\right) \tag{5.6.6}$$

as drift function and diffusion function, respectively. Note that σ_b^2 does not come into play in these functions.

If μ_c, σ_c^2 and σ_{bc} are of the same small enough order, then

$$\begin{aligned}\psi(y) = \exp\left(-2\int_0^y \frac{m(x)}{v(x)}dx\right) &\approx 1 - 2\int_0^y \left(-\mu_c + x(\sigma_{bc} - \sigma_c^2)\right)dx \\ &= 1 + 2\mu_c y - \left(\sigma_{bc} - \sigma_c^2\right)y^2 \end{aligned} \tag{5.6.7}$$

as in Lessard (2005). Therefore, the conditions

$$\int_0^1 \psi(y)dy < 1, \quad \psi(1) < 1, \quad \int_0^1 \psi(y)dy > \psi(1), \tag{5.6.8}$$

become

$$\sigma_{bc} - \sigma_c^2 > 3\mu_c, \quad \sigma_{bc} - \sigma_c^2 > 2\mu_c, \quad \sigma_{bc} - \sigma_c^2 > \frac{3}{2}\mu_c. \tag{5.6.9}$$

respectively. These are the conditions for selection to favor the evolution of C, favor more the evolution of C than the evolution of D, and disfavor the evolution of D, respectively. Since $\mu_c > 0$, these conditions can hold only if $\sigma_{bc} > \sigma_c^2$, in which case the first condition is the most stringent one and the third condition the least stringent one.

In the particular case $b(t) = rc(t)$ where $r > 0$ is a constant, the above conditions reduce to

$$r > 1 + 3\left(\frac{\mu_c}{\sigma_c^2}\right), \quad r > 1 + 2\left(\frac{\mu_c}{\sigma_c^2}\right), \quad r > 1 + \frac{3}{2}\left(\frac{\mu_c}{\sigma_c^2}\right), \tag{5.6.10}$$

respectively. These conditions can hold for $r > 1$ if σ_c^2 is large enough compared to μ_c. Moreover, it can be shown that at least the second condition does not depend on the assumption that σ_c^2 and σ_{bc} are small and of the same order (see **Appendix F**).

5.7 Repeated RPD

We turn now our attention to a RPD game that is repeated a random number of times. There are two pure actions, cooperation (C) and defection (D), and the payoffs in a single round of interaction between two players at time $t \geq 0$ are given by the random game matrix

$$\begin{pmatrix} R(t) & S(t) \\ T(t) & P(t) \end{pmatrix}. \tag{5.7.1}$$

Here, $R(t)$ and $S(t)$ are the payoffs to action C against C and D, respectively, while $T(t)$ and $P(t)$ are the corresponding payoffs to action D against the same two actions. These payoffs are assumed to be independent random variables whose distributions do not depend on time $t \geq 0$. Moreover, their expected values determine a classical PD game. Actually, we assume

$$\mathbf{E}(R(t)) = \mu_R N^{-1} > 0, \tag{5.7.2a}$$

$$\mathbf{E}(S(t)) = \mu_S N^{-1} > 0, \tag{5.7.2b}$$

$$\mathbf{E}(T(t)) = \mu_T N^{-1} > 0, \tag{5.7.2c}$$

$$\mathbf{E}(P(t)) = \mu_P N^{-1} > 0, \tag{5.7.2d}$$

with $\mu_T > \mu_R > \mu_P > \mu_S$ and $2\mu_R > \mu_T + \mu_S$. Finally, at each time $t \geq 0$, the number of rounds of interaction between the same two players is a random variable $n(t) \geq 1$ that is independent of $R(t)$, $S(t)$, $T(t)$ and $P(t)$.

In this repeated RPD game, we consider two strategies, Tit-for-Tat (TFT) and Always-Defect (AllD). In a pairwise interaction, a TFT-strategist uses action C in the first round and, in each of the next rounds, copies the action previously used by the opponent. On the other hand, an AllD-strategist uses action D in all the rounds. Thus, the payoffs to these two strategies at time $t \geq 0$ are given by

$$\begin{pmatrix} \eta_1(t) & \eta_2(t) \\ \eta_3(t) & \eta_4(t) \end{pmatrix} = \begin{pmatrix} n(t)R(t) & S(t) + (n(t)-1)P(t) \\ T(t) + (n(t)-1)P(t) & n(t)P(t) \end{pmatrix}. \tag{5.7.3}$$

Moreover, the population-scaled parameters (5.2.3) for the means, variances and covariances of the these payoffs take the form

$$\mu_1 = \mu_R \mathbf{E}(n(t)), \tag{5.7.4a}$$

$$\mu_2 = \mu_S + \mu_P \mathbf{E}(n(t) - 1), \tag{5.7.4b}$$

$$\mu_3 = \mu_T + \mu_P \mathbf{E}(n(t) - 1), \tag{5.7.4c}$$

$$\mu_4 = \mu_P \mathbf{E}(n(t)), \tag{5.7.4d}$$

and

$$\sigma_1^2 = \sigma_R^2 \mathbf{E}(n(t)^2), \tag{5.7.5a}$$

$$\sigma_2^2 = \sigma_S^2 + \sigma_P^2 \mathbf{E}((n(t) - 1)^2), \tag{5.7.5b}$$

$$\sigma_3^2 = \sigma_T^2 + \sigma_P^2 \mathbf{E}((n(t) - 1)^2), \tag{5.7.5c}$$

$$\sigma_4^2 = \sigma_P^2 \mathbf{E}(n(t)^2), \tag{5.7.5d}$$

$$\sigma_{23} = \sigma_P^2 \mathbf{E}((n(t) - 1)^2), \tag{5.7.5e}$$

$$\sigma_{24} = \sigma_{34} = \sigma_P^2 \mathbf{E}(n(t)(n(t) - 1)), \tag{5.7.5f}$$

$$\sigma_{12} = \sigma_{13} = \sigma_{14} = 0, \tag{5.7.5g}$$

where $\sigma_R^2 = N\mathbf{Var}(R(t))$, $\sigma_S^2 = N\mathbf{Var}(S(t))$, $\sigma_T^2 = N\mathbf{Var}(T(t))$ and $\sigma_P^2 = N\mathbf{Var}(P(t))$.

Substituting the above expressions into (5.3.4) and (5.3.5) yields

$$\begin{aligned} m(x) = x(1-x)\Big(&\mu_S - \mu_P + x(\mu_R - \mu_S - \mu_T + \mu_P) + x(\mu_R - \mu_P)\mathbf{E}(n(t) - 1) \\ &- x^3\sigma_R^2 \mathbf{E}(n(t)^2) - x(1-x)^2\sigma_S^2 + x^2(1-x)\sigma_T^2 \\ &+ (1-x)\sigma_P^2 \mathbf{E}\big(((n(t)-1)x + 1 - x)((n(t)-1)(1+x) + 1 - x)\big)\Big) \end{aligned} \tag{5.7.6}$$

and

$$\begin{aligned} v(x) = x(1-x)\Big(&1 + x^3(1-x)\sigma_R^2 \mathbf{E}(n(t)^2) + x(1-x)^3\sigma_S^2 \\ &+ x^3(1-x)\sigma_T^2 + x(1-x)\sigma_P^2 \mathbf{E}\big(((n(t)-1)x + 1 - x)^2\big)\Big) \end{aligned} \tag{5.7.7}$$

as drift function and diffusion function, respectively. Assuming that μ_R, μ_S, μ_T and μ_P, as well as σ_R^2, σ_S^2, σ_T^2 and σ_P^2, are of the same small enough order, we have

$$\psi(y) = \exp\left(-2\int_0^y \frac{m(x)}{v(x)}dx\right) \approx 1 - 2\int_0^y \frac{m(x)}{x(1-x)}dx \tag{5.7.8}$$

as in Lessard (2005). Then, the condition for the evolution of TFT to be favored by selection when introduced as a single mutant, which is given by

$$\int_0^1 \psi(y)dy < 1, \tag{5.7.9}$$

becomes

$$2\int_0^1 \int_0^y \frac{m(x)}{x(1-x)}dxdy > 0. \tag{5.7.10}$$

Using the expression of $m(x)$ given in (5.7.6), this condition can be written in the form

$$\begin{aligned}\mu_S - \mu_P &+ \frac{1}{3}(\mu_R - \mu_S - \mu_T + \mu_P) + \frac{1}{3}(\mu_R - \mu_P)\mathbf{E}(n(t) - 1) \\ &- \frac{1}{10}\mathbf{E}(n(t)^2)\sigma_R^2 - \frac{1}{10}\sigma_S^2 + \frac{1}{15}\sigma_T^2 + c_P\sigma_P^2 > 0,\end{aligned} \tag{5.7.11}$$

where

$$\begin{aligned}c_P &= 2\mathbf{E}((n(t)-1)^2)\int_0^1\int_0^y (1-x)x(1+x)dxdy \\ &\quad + 2\mathbf{E}(n(t)-1)\int_0^1\int_0^y (1-x)^2(1+2x)dxdy \\ &\quad + 2\int_0^1\int_0^y (1-x)^3dxdy \\ &= \frac{7}{30}\mathbf{E}((n(t)-1)^2) + \frac{7}{10}\mathbf{E}(n(t)-1) + \frac{2}{5} > 0.\end{aligned} \tag{5.7.12}$$

If all the variances of the payoffs vanish, then the condition (5.7.11) corresponds to the one-third law of evolution (Nowak *et al.*, 2004; Lessard, 2005), since it says then that the mean payoff to TFT exceeds the mean payoff to AllD when the frequency of TFT is equal to 1/3. Note that this condition holds if the expected number of rounds $\mathbf{E}(n(t))$ is large enough, since $\mu_R > \mu_P$. On the other hand, when the variances of the payoffs do not vanish, we see that an increase of σ_T^2 and σ_P^2, or a decrease of σ_R^2 and σ_S^2, makes it easier for the evolution of TFT to be favored by selection.

Similarly, from

$$\phi(y) = \exp\left(2\int_y^1 \frac{m(x)}{v(x)}dx\right) \approx 1 + 2\int_y^1 \frac{m(x)}{x(1-x)}dx, \tag{5.7.13}$$

the condition for the evolution of AllD not to be favored by selection when introduced as a single mutant, that is,

$$\int_0^1 \phi(y)dy > 1, \tag{5.7.14}$$

reduces to

$$2\int_0^1\int_y^1 \frac{m(x)}{x(1-x)}dxdy = 2\int_0^1\int_0^{1-y} \frac{m(1-x)}{x(1-x)}dxdy > 0, \tag{5.7.15}$$

which is equivalent to

$$\begin{aligned}\mu_S - \mu_P + \frac{2}{3}(\mu_R - \mu_S - \mu_T + \mu_P) + \frac{2}{3}(\mu_R - \mu_P)\mathbf{E}(n(t) - 1)\\ - \frac{2}{5}\mathbf{E}(n(t)^2)\sigma_R^2 - \frac{1}{15}\sigma_S^2 + \frac{1}{10}\sigma_T^2 + c_P\sigma_P^2 > 0,\end{aligned} \tag{5.7.16}$$

where

$$c_P = \frac{4}{15}\mathbf{E}((n(t)-1)^2) + \frac{3}{10}\mathbf{E}(n(t)-1) + \frac{1}{10} > 0. \tag{5.7.17}$$

Therefore, an increase of σ_T^2 and σ_P^2, or a decrease of σ_R^2 and σ_S^2, makes it easier also for the evolution of AllD not to be favored by selection.

Finally, we have

$$\psi(1) \approx 1 - 2\int_0^1 \frac{m(x)}{x(1-x)}dx < 1 \tag{5.7.18}$$

when

$$\begin{aligned}\int_0^1 \frac{m(x)}{x(1-x)}dx = \mu_S - \mu_P + \frac{1}{2}(\mu_R - \mu_S - \mu_T + \mu_P) + \frac{1}{2}(\mu_R - \mu_P)\mathbf{E}(n(t) - 1)\\ - \frac{1}{4}\mathbf{E}(n(t)^2)\sigma_R^2 - \frac{1}{12}\sigma_S^2 + \frac{1}{12}\sigma_T^2 + c_P\sigma_P^2 > 0,\end{aligned} \tag{5.7.19}$$

where

$$c_P = \frac{1}{4}\mathbf{E}((n(t)-1)^2) + \frac{1}{2}\mathbf{E}(n(t)-1) + \frac{1}{4} > 0. \qquad (5.7.20)$$

This means that an increase of σ_T^2 and σ_P^2, or a decrease of σ_R^2 and σ_S^2, makes it easier for selection to favor more the evolution of TFT than the evolution of AllD.

5.8 Discussion

Environmental noise in the payoffs of a matrix game may have important effects on the evolutionary dynamics, and even change the outcome of evolution. As a matter of fact, the dynamics is driven not only by the expected values of the payoffs but also by their variances. Variability in payoffs can push the time average of a population state far from its interior Nash equilibrium (Broom, 2005) or even change the stability of a fixation state (Stollmeier and Nagler, 2018). In the case of a deterministic one-round Prisoner's Dilemma (PD), where all the payoffs are constant, cooperation can never be favored by natural selection. However, introducing uncertainty in the payoffs makes it possible for cooperation to be favored.

Assuming a Randomized Prisoner's Dilemma (RPD) with independent payoffs in a large finite population, we have shown that, if the means and variances of the payoffs are of the same order of magnitude given by the inverse of the population size N, increasing the variance in the payoffs for defection, tends to promote the evolution of cooperation (conclusion 1). Moreover, if the payoffs have additive expected values, decreasing the variance in the payoffs for cooperation, at least for an expected cost for cooperation small enough, has the same effect (conclusion 2). More precisely, increasing the variance of the payoff for defection against defection (case 1) increases the probability of ultimate fixation of cooperation introduced as a single mutant, F_C, while increasing the variance of the payoff for defection against cooperation (case 2) decreases the probability of ultimate fixation of defection introduced as a single mutant, F_D. Increasing both variances simultaneously (cases 3 and 4) enhances the effect.

In particular, we have shown that the evolution of cooperation is fully favored by selection, in the sense that $F_C > N^{-1} > F_D$, where N^{-1} is the probability of ultimate fixation of a single mutant under neutrality, if the population-scaled variance of the payoffs for defection against cooperation and defection, σ^2, exceeds $(15c + 7\sigma_0^2)/3$, where σ_0^2 is the population-scaled variance of the payoffs for cooperation against cooperation and defection (case 4 and case 3 for $\sigma_0^2 = 0$). Moreover, as σ^2 is increased, the conditions for $F_C > N^{-1}$, $F_C > F_D$, and $F_D < N^{-1}$ are satisfied when $\sigma^2 > \sigma_*^2$, $\sigma^2 > \sigma_{**}^2$, and $\sigma^2 > \sigma_{***}^2$, respectively, where σ_*^2, σ_{**}^2 and σ_{***}^2 represent three increasing threshold values (Figure 2). These are the conditions for selection to favor the evolution of cooperation, favor more the evolution of cooperation than the evolution of defection, and disfavor the evolution of defection, respectively. We have analogous conditions with increasing threshold values when only the population-scaled variance of the payoff for defection against defection is increased (case 1), and with decreasing threshold values when only the population-scaled variance of the payoff for defection against cooperation is increased (case 2).

Our results are in agreement with the fact that, in the case of a RPD in an infinite population, a larger variance of the payoffs for defection is required for C-fixation to be stochastically locally stable and D-fixation stochastically locally unstable (Li *et al.*, 2019). On the other hand, they significantly differ from results obtained with constant payoffs in finite populations. For instance, in the case of a PD game in a graph-structured population and the case of a repeated PD game in a well-mixed population (Nowak *et al.*, 2004; Nowak, 2006), the condition $F_C > N^{-1}$ is sufficient for $F_C > N^{-1} > F_D$.

Note that more uncertainty in the payoffs for defection than for cooperation makes sense. Among the reasons, defectors are more isolated than cooperators and may not share with others the increments or decrements of surrounding resources caused by variations in the environment. They may also suffer from punishment or lack of reward from others. Even if the expected payoffs may still be higher for defection than for cooperation, their variances may also be higher.

On the other hand, when increasing the variance of all the payoffs (case 5), we have shown that selection can favor the evolution of both cooperation and defection in the sense that $F_C > N^{-1}$ and $F_D > N^{-1}$. Note that, since genotypic fitnesses in a random mating diploid population can be viewed as payoffs in random pairwise interactions of haploid individuals (in which case the payoff matrix is symmetric), our results extend previous results stated without proofs for population genetics models (Karlin and Levikson, 1974). Increasing the variance of the payoffs for defection (cooperation, respectively) pushes the system away from fixation of defection (cooperation, respectively), and at the same time promotes fixation of cooperation (defection, respectively). When the variance of all the payoffs increases, the system state is more likely to stay away from fixation.

Of further interest is the effect of the variances of the cost and benefit in a RPD with additive payoffs which are not independent. At least when the population-scaled means and variances of the cost and benefit, as well as their population-scaled covariance, are of the same small enough order, the conditions for $F_C > N^{-1}$, $F_C > F_D$ and $F_D < N^{-1}$ take the form $\sigma_{bc} - \sigma_c^2 > 3\mu_c$, $\sigma_{bc} - \sigma_c^2 > 2\mu_c$ and $\sigma_{bc} - \sigma_c^2 > 3\mu_c/2$, respectively, where μ_c is the population-scaled expected cost, σ_{bc} the population-scaled covariance between the cost and benefit, and σ_c^2 the population-scaled variance of the cost. The first condition is the most stringent one and the last condition the least stringent one, but they all require that $\sigma_{bc} > \sigma_c^2$ since $\mu_c > 0$. Of course, this does not occur if the cost and benefit are constants or independent random variables.

In the case of a repeated RPD game, the payoffs to TFT and AllD in pairwise interactions with a random number of rounds between the same players are generally not independent even if the payoffs to cooperation and defection are independent in each round. Assuming that the population-scaled means and variances of these payoffs are of the same small enough order, we have shown that an increase in the variances of the payoffs for defection, or a decrease in the variances of the payoffs for cooperation, makes it easier for $F_{TFT} > N^{-1}$, $F_{TFT} > F_{AllD}$ and $F_{AllD} < N^{-1}$ to hold. Since fixation of TFT means fixation of cooperation, the conclusion is that these conditions tend to promote the evolution of cooperation in agreement with our results for a one-round RPD game.

As a final remark, our results are based on a diffusion approximation that has been ascertained for a randomized matrix game with payoffs that have expected values, variances and covariances of order given by the inverse of a large population size N. This approximation can be used to study not only fixation probabilities, but any dynamical properties.

5.9 Appendix

5.9.1 A. Conditional moments of Δx

First fourth moments of a binomial distribution

Let $\tilde{x}$ be a random variable such that $N\tilde{x}$ follows a binomial distribution of parameters N and x, noted $\mathrm{B}(N, x)$. The first moment of $\tilde{x}$ is $\mathbf{E}(\tilde{x}) = x$. As for the second moment, we have

$$\begin{aligned}
\mathbf{E}(\tilde{x}^2) &= \sum_{i=0}^{N} \left(\frac{i}{N}\right)^2 \binom{N}{i} x^i (1-x)^{N-i} \\
&= \frac{1}{N^2}\left[\sum_{i=0}^{N} i(i-1)\binom{N}{i} x^i(1-x)^{N-i} + \sum_{i=0}^{N} i\binom{N}{i} x^i (1-x)^{N-i}\right] \\
&= \frac{1}{N^2}\left[N(N-1)x^2 \sum_{i=2}^{N} \binom{N-2}{i-2} x^{i-2}(1-x)^{N-i}\right] + \frac{x}{N} \\
&= \frac{N-1}{N}x^2 + \frac{x}{N} \\
&= x^2 + \frac{x(1-x)}{N}. \qquad (5.9.1)
\end{aligned}$$

Analogously, using the above expression for $\mathbf{E}(\tilde{x}^2)$, the third moment is given by

$$\begin{aligned}
\mathbf{E}(\tilde{x}^3) &= \sum_{i=0}^{N}\left(\frac{i}{N}\right)^3\binom{N}{i}x^i(1-x)^{N-i} \\
&= \frac{x(N-1)^2}{N^2}\left[\sum_{i=1}^{N}\frac{(i-1)^2+2(i-1)+1}{(N-1)^2}\binom{N-1}{i-1}x^{i-1}(1-x)^{(N-1)-(i-1)}\right] \\
&= \frac{x(N-1)^2}{N^2}\left[x^2+\frac{x(1-x)}{N}+\frac{2}{N-1}x+\frac{1}{(N-1)^2}\right] \\
&= x^3+\frac{3x^2(1-x)}{N}+o(N^{-1}). \qquad (5.9.2)
\end{aligned}$$

Finally, as for the fourth moment, we find

$$\begin{aligned}
\mathbf{E}(\tilde{x}^4) &= \sum_{i=0}^{N}\left(\frac{i}{N}\right)^4\binom{N}{i}x^i(1-x)^{N-i} \\
&= \frac{x(N-1)^3}{N^3} \\
&\times\left[\sum_{i=1}^{N}\frac{(i-1)^3+3(i-1)^2+3(i-1)+1}{(N-1)^3}\binom{N-1}{i-1}x^{i-1}(1-x)^{(N-1)-(i-1)}\right] \\
&= \frac{x(N-1)^3}{N^3} \\
&\times\left[x^3+\frac{3x^2(1-x)}{N}+\frac{3}{N-1}(x^2+\frac{x(1-x)}{N})+\frac{3}{(N-1)^2}x+\frac{1}{(N-1)^3}+o(N^{-1})\right] \\
&= x^4+\frac{6x^3(1-x)}{N}+o(N^{-1}). \qquad (5.9.3)
\end{aligned}$$

First conditional moments of Δx

Given that $x(t)=x$, the frequency change $\Delta x = x(t+\Delta t)-x(t)$ has the same probability distribution as $\tilde{x}-x$, where $N\tilde{x}\sim \mathrm{B}(N,x')$. Here, the parameter

$$x' = \frac{x\mathbf{\Pi}_1}{x\mathbf{\Pi}_1+(1-x)\mathbf{\Pi}_2} \qquad (5.9.4)$$

is a random variable with

$$\mathbf{\Pi}_1 = 1+x\eta_1+(1-x)\eta_2 \qquad (5.9.5)$$

and

$$\mathbf{\Pi}_2 = 1+x\eta_3+(1-x)\eta_4, \qquad (5.9.6)$$

where η_i has mean $N^{-1}\mu_i$, variance $N^{-1}\sigma_i^2 + o(N^{-1})$ and covariance $N^{-1}\sigma_{ij} + o(N^{-1})$ with η_j for $i, j = 1, \ldots, 4$ with $j \neq i$. Note that

$$x' = \frac{x(1 + P_1)}{1 + P_3}, \tag{5.9.7}$$

where

$$P_1 = x\eta_1 + (1 - x)\eta_2, \tag{5.9.8a}$$

$$P_2 = x\eta_3 + (1 - x)\eta_4, \tag{5.9.8b}$$

$$P_3 = xP_1 + (1 - x)P_2. \tag{5.9.8c}$$

The random variables P_j for $j = 1, 2$ are homogeneous linear functions of $\eta_1, \ldots, \eta_4$, while P_3 is a homogeneous linear function of P_1 and P_2. Thus, the moments of P_1 and P_2 satisfy

$$\mathbf{E}(P_1) = N^{-1}\left(x\mu_1 + (1 - x)\mu_2\right), \tag{5.9.9a}$$

$$\mathbf{E}(P_2) = N^{-1}\left(x\mu_3 + (1 - x)\mu_4\right), \tag{5.9.9b}$$

$$\mathbf{E}(P_1^2) = N^{-1}\left(x^2\sigma_1^2 + (1 - x)^2\sigma_2^2 + 2x(1 - x)\sigma_{12}\right) + o(N^{-1}), \tag{5.9.9c}$$

$$\mathbf{E}(P_2^2) = N^{-1}\left(x^2\sigma_3^2 + (1 - x)^2\sigma_4^2 + 2x(1 - x)\sigma_{34}\right) + o(N^{-1}), \tag{5.9.9d}$$

$$\mathbf{E}(P_1P_2) = N^{-1}\left(x^2\sigma_{13} + x(1 - x)(\sigma_{14} + \sigma_{23}) + (1 - x)^2\sigma_{24}\right) + o(N^{-1}), \tag{5.9.9e}$$

and

$$\mathbf{E}(P_1^k P_2^l) = o(N^{-1}) \tag{5.9.10}$$

as soon as k, l are non-negative integers such that $k + l \geq 3$. Therefore, we have

$$\begin{aligned}
\mathbf{E}(x') &= \mathbf{E}\big(x(1+P_1)(1-P_3+P_3^2)\big) + o(N^{-1}) \\
&= \mathbf{E}\big(x + x(1-x)(P_1-P_2) + x(1-x)(P_2-P_1)(xP_1+(1-x)P_2)\big) + o(N^{-1}) \\
&= x + x(1-x)\big(\mathbf{E}(P_1) - \mathbf{E}(P_2)\big) \\
&\quad + x(1-x)\big(-x\mathbf{E}(P_1^2) + (1-x)\mathbf{E}(P_2^2) + (2x-1)\mathbf{E}(P_1P_2)\big) + o(N^{-1}) \\
&= x + \frac{x(1-x)}{N}\big(\mu_2 - \mu_4 + x(\mu_1 - \mu_2 - \mu_3 + \mu_4)\big) \\
&\quad + \frac{x(1-x)}{N}\Big(- x^3\sigma_1^2 - x(1-x)^2\sigma_2^2 - 2x^2(1-x)\sigma_{12} \\
&\quad + x^2(1-x)\sigma_3^2 + (1-x)^3\sigma_4^2 + 2x(1-x)^2\sigma_{34} \\
&\quad + (2x-1)\left(x^2\sigma_{13} + x(1-x)(\sigma_{14}+\sigma_{23}) + (1-x)^2\sigma_{24}\right)\Big) + o(N^{-1}). \qquad (5.9.11)
\end{aligned}$$

Since

$$\mathbf{E}(\Delta x | x(t) = x) = \mathbf{E}(\tilde{x} - x) = \mathbf{E}(\tilde{x}) - x = \mathbf{E}(x') - x, \qquad (5.9.12)$$

the first conditional moment of Δx is given by

$$\begin{aligned}
\mathbf{E}(\Delta x | x(t) = x) = \frac{x(1-x)}{N}\Big(&\mu_2 - \mu_4 + x\big(\mu_1 - \mu_2 - \mu_3 + \mu_4\big) \\
&+ x^3(\sigma_{13} - \sigma_1^2) + x(1-x)^2(2\sigma_{34} - \sigma_{14} - \sigma_{23} + \sigma_{24} - \sigma_2^2) \\
&+ x^2(1-x)(-2\sigma_{12} + \sigma_{14} + \sigma_{23} - \sigma_{13} + \sigma_3^2) + (1-x)^3(\sigma_4^2 - \sigma_{24})\Big) \\
&+ o(N^{-1}), \qquad (5.9.13)
\end{aligned}$$

while its second conditional moment can be expressed as

$$\begin{aligned}
\mathbf{E}((\Delta x)^2 | x(t) = x) &= \mathbf{E}\big((\tilde{x} - x)^2\big) \\
&= \mathbf{E}\big(\tilde{x}^2 - 2x\tilde{x} + x^2\big) \\
&= \mathbf{E}\left(x'^2 + \frac{x'(1-x')}{N}\right) - 2x\mathbf{E}(x') + x^2 \\
&= \mathbf{E}\left((x'-x)^2\right) + \mathbf{E}\left(\frac{x'(1-x')}{N}\right) \\
&= \mathbf{E}\left((x'-x)^2\right) + \frac{x(1-x)}{N} + o(N^{-1}). \qquad (5.9.14)
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
\mathbf{E}\left((x'-x)^2\right) &= \mathbf{E}\Big(\big(x(1-x)(P_2-P_1)(P_3-1)\big)^2\Big) + o(N^{-1}) \\
&= \mathbf{E}\big(x^2(1-x)^2(P_2-P_1)^2\big) + o(N^{-1}) \\
&= \frac{x^2(1-x)^2}{N}\Big(x^2(\sigma_1^2+\sigma_3^2) + (1-x)^2(\sigma_2^2+\sigma_4^2) + 2x(1-x)(\sigma_{12}+\sigma_{34}) \\
&\quad - 2x^2\sigma_{13} - 2x(1-x)(\sigma_{14}+\sigma_{23}) - 2(1-x)^2\sigma_{24}\Big) + o(N^{-1}). \qquad (5.9.15)
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
&\mathbf{E}((\Delta x)^2|x(t)=x) \\
&= \frac{x(1-x)}{N}\big(1 + x^3(1-x)(\sigma_1^2+\sigma_3^2-2\sigma_{13}) + x(1-x)^3(\sigma_2^2+\sigma_4^2-2\sigma_{24}) \\
&\quad + 2x^2(1-x)^2(\sigma_{12}+\sigma_{34}-\sigma_{14}-\sigma_{23})\big) + o(N^{-1}). \qquad (5.9.16)
\end{aligned}$$

Finally, the fourth conditional moment of Δx can be expressed as

$$\begin{aligned}
\mathbf{E}((\Delta x)^4|x(t)=x) &= \mathbf{E}\big((\tilde{x}-x)^4\big) \\
&= \mathbf{E}\big(\tilde{x}^4 - 4x\tilde{x}^3 + 6x^2\tilde{x}^2 - 4x^3\tilde{x} + x^4\big) \\
&= \mathbf{E}\left(x'^4 + \frac{6x'^3(1-x')}{N}\right) - 4x\mathbf{E}\left(x'^3 + \frac{3x'^2(1-x')}{N}\right) \\
&\quad + 6x^2\mathbf{E}\left(x'^2 + \frac{x'(1-x')}{N}\right) - 4x^3\mathbf{E}(x') + x^4 + o(N^{-1}) \\
&= \mathbf{E}\left((x'-x)^4\right) + \frac{6}{N}\mathbf{E}\left(x'(1-x')(x'-x)^2\right) + o(N^{-1}). \qquad (5.9.17)
\end{aligned}$$

Note that

$$\mathbf{E}\left((x'-x)^4\right) = \mathbf{E}\big(x^4(1-x)^4(P_2-P_1)^4\big) + o(N^{-1}) = o(N^{-1}) \qquad (5.9.18)$$

and

$$\mathbf{E}\left(x'(1-x')(x'-x)^2\right) = \mathbf{E}\big(x'(1-x')x^2(1-x)^2(P_2-P_1)^2\big) + o(N^{-1}) = O(N^{-1}), \qquad (5.9.19)$$

from which we conclude that $\mathbf{E}((\Delta x)^4|x(t)=x) = o(N^{-1})$.

5.9.2 B. Partial derivatives of $g(x)$ with respect to σ_i^2

The function $g(x)$ is defined as the drift function in (5.3.4) divided by the diffusion function in (5.3.5), that is,

$$g(x) = \frac{m(x)}{v(x)} = \frac{h(x) - \left(x^3\sigma_1^2 + x(1-x)^2\sigma_2^2 - x^2(1-x)\sigma_3^2 - (1-x)^3\sigma_4^2\right)}{1 + x^3(1-x)(\sigma_1^2+\sigma_3^2) + x(1-x)^3(\sigma_2^2+\sigma_4^2)}, \tag{5.9.20}$$

where

$$h(x) = \mu_2 - \mu_4 + x(\mu_1 - \mu_2 - \mu_3 + \mu_4). \tag{5.9.21}$$

Then the partial derivatives with respect to $\sigma_i^2 > 0$ for $i = 1, \ldots, 4$ are given by

$$\frac{\partial g(x)}{\partial \sigma_1^2} = -\frac{x^3\left(1 + (1-x)h(x) + x^2(1-x)\sigma_3^2 + (1-x)^3\sigma_4^2\right)}{\left(1 + x^3(1-x)(\sigma_1^2+\sigma_3^2) + x(1-x)^3(\sigma_2^2+\sigma_4^2)\right)^2}, \tag{5.9.22a}$$

$$\frac{\partial g(x)}{\partial \sigma_2^2} = -\frac{x(1-x)^2\left(1 + (1-x)h(x) + x^2(1-x)\sigma_3^2 + (1-x)^3\sigma_4^2\right)}{\left(1 + x^3(1-x)(\sigma_1^2+\sigma_3^2) + x(1-x)^3(\sigma_2^2+\sigma_4^2)\right)^2}, \tag{5.9.22b}$$

$$\frac{\partial g(x)}{\partial \sigma_3^2} = \frac{x^2(1-x)\left(1 - xh(x) + x^3\sigma_1^2 + x(1-x)^2\sigma_2^2\right)}{\left(1 + x^3(1-x)(\sigma_1^2+\sigma_3^2) + x(1-x)^3(\sigma_2^2+\sigma_4^2)\right)^2}, \tag{5.9.22c}$$

$$\frac{\partial g(x)}{\partial \sigma_4^2} = \frac{(1-x)^3\left(1 - xh(x) + x^3\sigma_1^2 + x(1-x)^2\sigma_2^2\right)}{\left(1 + x^3(1-x)(\sigma_1^2+\sigma_3^2) + x(1-x)^3(\sigma_2^2+\sigma_4^2)\right)^2}. \tag{5.9.22d}$$

Under the assumptions $\mu_4 > \mu_2$ and $\mu_3 > \mu_1$, we have $h(x) < 0$ for $0 < x < 1$, in which case

$$\frac{\partial g(x)}{\partial \sigma_3^2} > 0 \tag{5.9.23}$$

and

$$\frac{\partial g(x)}{\partial \sigma_4^2} > 0. \tag{5.9.24}$$

On the other hand, the conditions

$$\frac{\partial g(x)}{\partial \sigma_1^2} < 0 \tag{5.9.25}$$

and

$$\frac{\partial g(x)}{\partial \sigma_2^2} < 0 \tag{5.9.26}$$

hold if and only if

$$h(x) > -\frac{1 + x^2(1-x)\sigma_3^2 + (1-x)^3\sigma_4^2}{1-x}. \tag{5.9.27}$$

Since the right-hand member in (5.9.27) is less than -1 for $0 < x < 1$, a sufficient condition for (5.9.25) and (5.9.25) to hold is $h(x) \geq -1$ on $(0, 1)$.

5.9.3 C. Monotonicity of $g(x)$ and convexity of $\psi(y)$, $\phi(y)$ in cases 1, 3-5

In case 1, with $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 0$ and $\sigma_4^2 = \sigma^2 > 0$, the expression of $g(x)$ takes the form

$$g(x) = \frac{-c + (1-x)^3\sigma^2}{1 + x(1-x)^3\sigma^2}, \tag{5.9.28}$$

whose derivative is given by

$$g'(x) = \frac{\sigma^2(1-x)^2\left[(c-\sigma^2)(1-x)^4 - cx^2(2+(2-x)^2) - 3\right]}{\left(1 + x(1-x)^3\sigma^2\right)^2}. \tag{5.9.29}$$

If $\sigma^2 > c$, then $g'(x) < 0$ for $0 < x < 1$. Thus, $g(x)$ is a strictly decreasing function on $[0, 1]$.

In case 3, with $\sigma_1^2 = \sigma_2^2 = 0$ and $\sigma_3^2 = \sigma_4^2 = \sigma^2 > 0$, we have

$$g(x) = \frac{-c + (1-x)a\sigma^2}{1 + x(1-x)a\sigma^2}, \tag{5.9.30}$$

where $a = x^2 + (1-x)^2$. Under the condition $\sigma^2 > c$, the derivative of $g(x)$ for $0 < x < 1$ is given by

$$g'(x) = \frac{2x - 3a + c\left(2x^2 + (1-4x)a\right) - \sigma^2 a^2 (1-x)^2}{\left(1 + x(1-x)a\sigma^2\right)^2}, \tag{5.9.31}$$

where $2x - 3a = -6x^2 + 8x - 3 \leq -\frac{1}{3} < 0$ and

$$\begin{aligned} c\left(2x^2 + (1-4x)a\right) - \sigma^2 a^2(1-x)^2 &< c\left(2x^2 + (1-4x)a\right) - ca^2(1-x)^2 \\ &= c\left(a(1 - 4x - (1-x)^4 - x^2(1-x)^2) + 2x^2\right) \\ &= cx^2\left(a(-7 + 6x - 2x^2) + 2\right) \\ &= cx^2\left(-4x^4 + 16x^3 - 28x^2 + 20x - 5\right) \\ &= -cx^2\left(4(x-1)^4 + (2x-1)^2\right) \\ &< 0. \end{aligned} \tag{5.9.32}$$

Therefore, $g'(x) < 0$ for $0 < x < 1$. Consequently, the function $g(x)$ is strictly decreasing on $[0, 1]$.

In case 4, with $\sigma_1^2 = \sigma_2^2 = \sigma_0^2 > 0$, $\sigma_3^2 = \sigma_4^2 = \sigma^2 > 0$ and $a = x^2 + (1-x)^2$, we have

$$g(x) = \frac{-c + (1-x)a\sigma^2 - xa\sigma_0^2}{1 + x(1-x)a(\sigma^2 + \sigma_0^2)}. \tag{5.9.33}$$

Under the condition $\sigma^2 > c$, the derivative of $g(x)$ for $0 < x < 1$ is given by

$$\begin{aligned} g'(x) = {} & \frac{1}{\big(1 + x(1-x)a(\sigma^2+\sigma_0^2)\big)^2}\Big[(-6x^2 + 8x - 3)\sigma^2 + (-6x^2 - 4x - 1)\sigma_0^2 \\ & + (\sigma^2 + \sigma_0^2)\left(-c(2x-1)^3 - \sigma^2 a^2 (1-x)^2 - \sigma_0^2 a^2 x^2\right)\Big], \end{aligned} \tag{5.9.34}$$

where $-6x^2 + 8x - 3 \leq -\frac{1}{3} < 0$, $-6x^2 - 4x - 1 < 0$ and

$$\begin{aligned} -c(2x-1)^3 - \sigma^2 a^2 (1-x)^2 - \sigma_0^2 a^2 x^2 &< -c\left((2x-1)^3 + a^2(1-x)^2\right) - \sigma_0^2 a^2 x^2 \\ &= -cx^2\left(4(x-1)^4 + (2x-1)^2\right) - \sigma_0^2 a^2 x^2 \\ &< 0. \end{aligned} \tag{5.9.35}$$

Thus, $g'(x) < 0$ for $0 < x < 1$, from which the function $g(x)$ is strictly decreasing on $[0, 1]$.

Finally, in case 5, with $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma^2 > c > 0$ and $a = x^2 + (1-x)^2$, the expression of $g(x)$ takes the form

$$g(x) = \frac{-c + (1-2x)a\sigma^2}{1 + 2x(1-x)a\sigma^2}, \tag{5.9.36}$$

whose derivative for $0 < x < 1$ is

$$g'(x) = \frac{(-12x^2 + 12x - 4)\sigma^2 + 2\sigma^2\left(c(1-2x)^3 - \sigma^2 a^3\right)}{\left(1 + 2x(1-x)a\sigma^2\right)^2}, \tag{5.9.37}$$

where $-12x^2 + 12x - 4 \leq -1 < 0$ and

$$\begin{aligned} c(1-2x)^3 - \sigma^2 a^3 &< -c\left((2x-1)^3 + a^3\right) \\ &= -2cx^2\left((2x-1)^2 + 2a(x-1)^2\right) \\ &< 0. \end{aligned} \tag{5.9.38}$$

We conclude that $g'(x) < 0$ for $0 < x < 1$, from which the function $g(x)$ is strictly decreasing on $[0, 1]$.

In addition, with $\psi(y) = \exp\left(-2\int_0^y g(x)dx\right) > 0$, $\phi(y) = \exp\left(2\int_y^1 g(x)dx\right) > 0$ and $g'(y) < 0$ for $0 < y < 1$ in cases 1, 3, 4, 5, we have

$$\psi''(y) = 2\psi(y)\left(2g^2(y) - g'(y)\right) > 0 \tag{5.9.39}$$

and

$$\phi''(y) = 2\phi(y)\left(2g^2(y) - g'(y)\right) > 0 \tag{5.9.40}$$

on $(0, 1)$, from which $\psi(y)$ and $\phi(y)$ are strictly convex functions on [0,1].

5.9.4 D. Monotonicity of $\psi(y)$ with respect to σ^2 in case 5

In case 5, with $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma^2 > 0$, we have

$$g(x) = \frac{-c + (1-2x)a\sigma^2}{1 + 2x(1-x)a\sigma^2}, \tag{5.9.41}$$

where $a = x^2 + (1-x)^2$. We show that, with respect to $\sigma^2 > 0$ and for every $y \in (0, 1]$, the function

$$k(y) = -\frac{\ln\psi(y)}{2} = \int_0^y g(x)dx \tag{5.9.42}$$

is strictly increasing so that

$$\psi(y) = \exp\left(-2k(y)\right) \tag{5.9.43}$$

is strictly decreasing. For $y \in (0, 1)$, we find

$$\frac{\partial k(y)}{\partial\sigma^2} = \int_0^y \frac{\partial g(x)}{\partial\sigma^2}dx = \int_0^y \frac{a\left(2cx(1-x) - (2x-1)\right)}{\left(1 + 2ax(1-x)\sigma^2\right)^2}dx. \tag{5.9.44}$$

It is easy to check that the integrand in (5.9.44) is positive on $[0, x^*)$ and negative on $(x^*, 1]$ with

$$x^* = \frac{1}{2} + \frac{\sqrt{c^2+1} - 1}{2c}. \tag{5.9.45}$$

This implies that the derivative in (5.9.44) is strictly increasing on $[0, x^*)$ and strictly decreasing on $(x^*, 1]$. Moreover, $\partial k(0)/\partial\sigma^2 = 0$ and

$$\begin{aligned}
\frac{\partial k(1)}{\partial\sigma^2} &= \int_0^1 \frac{a\left(2cx(1-x)-(2x-1)\right)}{\left(1+2ax(1-x)\sigma^2\right)^2}dx \\
&= \int_0^{\frac{1}{2}} \frac{a\left(2cx(1-x)-(2x-1)\right)}{\left(1+2ax(1-x)\sigma^2\right)^2}dx + \int_{\frac{1}{2}}^1 \frac{a\left(2cx(1-x)-(2x-1)\right)}{\left(1+2ax(1-x)\sigma^2\right)^2}dx \\
&= \int_0^{\frac{1}{2}} \frac{a\left(2cx(1-x)-(2x-1)\right)}{\left(1+2ax(1-x)\sigma^2\right)^2}dx + \int_0^{\frac{1}{2}} \frac{a\left(2cx(1-x)+(2x-1)\right)}{\left(1+2ax(1-x)\sigma^2\right)^2}dx \\
&= \int_0^{\frac{1}{2}} \frac{4acx(1-x)}{\left(1+2ax(1-x)\sigma^2\right)^2}dx > 0. \qquad (5.9.46)
\end{aligned}$$

We can conclude that $\partial k(y)/\partial\sigma^2 > 0$ for $y \in (0, 1]$.

5.9.5 E. Existence of σ_*^2, σ_{**}^2 and σ_{***}^2 in cases 1 to 5

In cases 1 to 4, it is easy to check that

$$\lim_{\sigma^2\to\infty} g(x) = x^{-1}. \qquad (5.9.47)$$

From (5.4.13) and (5.4.14) shown in Appendix B, $g(x)$ for $0 < x < 1$ strictly increases with respect to $\sigma^2 > 0$ and, therefore, $\psi(1) = \exp\left(-2\int_0^1 g(x)dx\right)$ strictly decreases. Moreover, when $\sigma^2 = 0$, we have $g(x) < 0$ for every $x \in [0, 1]$. This implies that $\psi(1) > 1$ for $\sigma^2 > 0$. On the other hand, owing to Fatou's lemma, we have

$$\lim_{\sigma^2\to\infty} \int_0^1 g(x)dx = \liminf_{\sigma^2\to\infty} \int_0^1 g(x)dx \geq \int_0^1 \lim_{\sigma^2\to\infty} g(x)dx. \qquad (5.9.48)$$

Therefore, we get

$$\begin{aligned}
\lim_{\sigma^2\to\infty} \psi(1) &= \lim_{\sigma^2\to\infty} \exp\left(-2\int_0^1 g(x)dx\right) \\
&\leq \exp\left(-2\int_0^1 \lim_{\sigma^2\to\infty} g(x)dx\right) \\
&= \exp\left(-2\int_0^1 x^{-1}dx\right) \\
&= 0. \qquad (5.9.49)
\end{aligned}$$

We conclude that the equation $\psi(1) = 1$ has a unique solution with respect to $\sigma^2 > 0$, denoted by σ^2_{**}. Then, we have $F_C > F_D$ if and only if $\sigma^2 > \sigma^2_{**}$.

In cases 1, 3 and 4, where $\psi(y)$ is a strictly convex function, the conditions $\psi(0) = 1$ and $\psi(1) = 1$ when $\sigma^2 = \sigma^2_{**}$ entails $\int_0^1 \psi(y)dy < 1$ when $\sigma^2 = \sigma^2_{**}$. Since this integral strictly decreases with respect to σ^2 and $\int_0^1 \psi(y)dy > 1$ when $\sigma^2 = 0$, the equation $\int_0^1 \psi(y)dy = 1$ has a unique solution with respect to σ^2 between 0 and σ^2_{**}, denoted by σ^2_*. Then, we have $F_C > N^{-1}$ if and only if $\sigma^2 > \sigma^2_*$.

Moreover, we know that $F_D = F_C > N^{-1}$ when $\sigma^2 = \sigma^2_{**}$. Using Fatou's lemma and the inequality $e^x \geq 1 + x > x$ for $x > 0$, we get

$$\begin{aligned}
\lim_{\sigma^2\to\infty} F_D &= \lim_{\sigma^2\to\infty} \frac{1}{N\int_0^1 \phi(y)dy} \\
&= \left(N \lim_{\sigma^2\to\infty} \int_0^1 \exp\left(2\int_y^1 g(x)dx\right) dy\right)^{-1} \\
&\leq \left(N \int_0^1 \exp\left(2\int_y^1 \lim_{\sigma^2\to\infty} g(x)dx\right) dy\right)^{-1} \\
&= \left(N \int_0^1 \exp\left(2(y^{-2} - 1)\right) dy\right)^{-1} \\
&\leq \left(2N \int_0^1 (y^{-2} - 1)dy\right)^{-1} \\
&= 0. \qquad (5.9.50)
\end{aligned}$$

Owing to (5.4.13) and (5.4.14) proved in Appendix B, we know that $\phi(y) = \exp\left(2\int_y^1 g(x)dx\right)$ is strictly increasing with respect to $\sigma^2 > 0$ for $y \in [0, 1)$. Therefore, F_D is a strictly decreasing function of σ^2. Thus, there must exist a threshold value of $\sigma^2 > \sigma^2_{**}$, denoted by σ^2_{***}, which is the unique solution of the equation $\int_0^1 \phi(y)dy = 1$. If $\sigma^2 > \sigma^2_{***}$, then $F_D < N^{-1}$.

In case 2, it is still possible to ascertain the existence of σ^2_*, since $\int_0^1 \psi(y)dy > 1$ when $\sigma^2 = 0$ and

$$\lim_{\sigma^2\to\infty} \int_0^1 \psi(y)dy \leq \int_0^1 \exp\left(-2\int_0^y x^{-1}dx\right) dy = 0. \qquad (5.9.51)$$

However, $c < \sigma_*^2 < \sigma_{**}^2$ can no longer be guaranteed.

In case 5, we have

$$g(x) = \frac{-c + (1-2x)a\sigma^2}{1 + 2x(1-x)a\sigma^2} \tag{5.9.52}$$

with $a = x^2 + (1-x)^2$. Moreover, we find

$$\begin{aligned}
\int_0^1 g(x)dx &= \int_0^{\frac{1}{2}} g(x)dx + \int_{\frac{1}{2}}^1 g(x)dx \\
&= \int_0^{\frac{1}{2}} g(x)dx + \int_0^{\frac{1}{2}} g(1-x)dx \\
&= \int_0^{\frac{1}{2}} \frac{-c + (1-2x)a\sigma^2}{1 + 2x(1-x)a\sigma^2}dx + \int_0^{\frac{1}{2}} \frac{-c + (2x-1)a\sigma^2}{1 + 2x(1-x)a\sigma^2}dx \\
&= -2c \int_0^{\frac{1}{2}} \frac{1}{1 + 2x(1-x)a\sigma^2}dx. \qquad (5.9.53)
\end{aligned}$$

Since the integrand in (5.9.53) is bounded by 1 and uniformly converges to 0 as $\sigma^2 \to \infty$ on $[\varepsilon, 1/2]$ for $\varepsilon > 0$, we have

$$\begin{aligned}
\lim_{\sigma^2 \to \infty} \psi(1, \sigma^2) &= \lim_{\sigma^2 \to \infty} \exp\left(-2\int_0^1 g(x)dx\right) \\
&= \exp\left(-2 \lim_{\sigma^2 \to \infty} \int_0^1 g(x)dx\right) \\
&= \exp\left(4c \lim_{\sigma^2 \to \infty} \int_0^{\frac{1}{2}} \frac{1}{1 + 2x(1-x)a\sigma^2}dx\right) \\
&= 1. \qquad (5.9.54)
\end{aligned}$$

From Appendix C, we already know that $\psi(1)$ is strictly decreasing with respect to σ^2 for $y \in (0, 1]$. Thus, we have $\psi(1) > 1$ for $\sigma^2 > 0$, which implies that equation (5.4.11) can never be satisfied. As a matter of fact, we always have $F_C < F_D$ and σ_{**}^2 does not exist. Moreover, since $\phi(y)$ is a strictly convex function on $[0, 1]$ with $\phi(0) = \psi(1)^{-1} < 1$ and $\phi(1) = 1$, we have $\phi(y) < 1$ for $y \in [0, 1)$. This tells us that condition (5.4.10) can never be satisfied. We always have $F_D > N^{-1}$ and σ_{***}^2 does not exist. The only threshold value of σ^2 in this case is σ_*^2, since $\psi(y)$ is a strictly convex function on $[0, 1]$ and a strictly decreasing function with respect to σ^2 for $y \in (0, 1]$ (see Appendix D). Its boundary value

$\psi(0) = 1$ with $\psi'(0) = 2c - 2\sigma^2 < 0$ for σ^2 large enough, along with equation (5.9.54), guarantees the existence of σ_*^2 which is the unique value of $\sigma^2 > 0$ such that $\int_0^1 \psi(y)dy = 1$.

5.9.6 F. Condition for $\psi(1) < 1$ in the additive RPD with $b(t) = rc(t)$

We consider an additive RPD game where the benefit $b(t)$ is linear with respect to the cost $c(t)$, that is, $b(t) = rc(t)$. Here, r is a constant that represents the "benefit to cost ratio". Then we have $\sigma_{bc} = r\sigma_c^2$, from which

$$g(x) = \frac{-\mu_c + x(r-1)\sigma_c^2}{1 + x(1-x)\sigma_c^2}. \tag{5.9.55}$$

Moreover, the denominator can be expressed as

$$\begin{aligned} 1 + x(1-x)\sigma_c^2 &= 1 + \frac{\sigma_c^2}{4} - \left(-\frac{\sigma_c^2}{4} + x\sigma_c^2 - x^2\sigma_c^2\right) \\ &= \left(\frac{A}{2}\right)^2 - \left(\left(\frac{1}{2} - x\right)\sigma_c\right)^2 \\ &= \left(\frac{A}{2} + \left(\frac{1}{2} - x\right)\sigma_c\right)\left(\frac{A}{2} - \left(\frac{1}{2} - x\right)\sigma_c\right), \end{aligned} \tag{5.9.56}$$

where $A = \sqrt{4 + \sigma_c^2}$. Assuming $g(x)$ in the form

$$g(x) = \frac{S_1}{\frac{A}{2} + \left(\frac{1}{2} - x\right)\sigma_c} - \frac{S_2}{\frac{A}{2} - \left(\frac{1}{2} - x\right)\sigma_c}, \tag{5.9.57}$$

we get the equations

$$S_1\left(\frac{A}{2} - \frac{\sigma_c}{2}\right) - S_2\left(\frac{A}{2} + \frac{\sigma_c}{2}\right) = -\mu_c, \tag{5.9.58}$$

$$(S_1 + S_2)\sigma_c = (r-1)\sigma_c^2, \tag{5.9.59}$$

which are equivalent to

$$S_1 - S_2 = \frac{-2\mu_c + (r-1)\sigma_c^2}{A}, \tag{5.9.60}$$

$$S_1 + S_2 = (r-1)\sigma_c. \tag{5.9.61}$$

Thus, we get

$$\begin{aligned}
\psi(y) &= \exp\left(-2\int_0^y g(x)dx\right) \\
&= \exp\Big(-2\Big(-\frac{S_1}{\sigma_c}\ln\left(\frac{A}{2}+\left(\frac{1}{2}-y\right)\sigma_c\right)-\frac{S_2}{\sigma_c}\ln\left(\frac{A}{2}-\left(\frac{1}{2}-y\right)\sigma_c\right) \\
&\quad +\frac{S_1}{\sigma_c}\ln\left(\frac{A}{2}+\frac{\sigma_c}{2}\right)+\frac{S_2}{\sigma_c}\ln\left(\frac{A}{2}-\frac{\sigma_c}{2}\right)\Big)\Big) \\
&= \left(\frac{A+\sigma_c-2\sigma_c y}{A+\sigma_c}\right)^{\frac{S_1}{\sigma_c}}\left(\frac{A-\sigma_c+2\sigma_c y}{A-\sigma_c}\right)^{\frac{S_2}{\sigma_c}}, \qquad (5.9.62)
\end{aligned}$$

from which

$$\psi(1)=\left(\frac{A-\sigma_c}{A+\sigma_c}\right)^{\frac{S_1-S_2}{\sigma_c}}. \qquad (5.9.63)$$

Since $A+\sigma_c \geq A-\sigma_c > 0$, the condition $\psi(1) < 1$ is satisfied if and only if $S_1 - S_2 > 0$, which means

$$r-1 > \frac{2\mu_c}{\sigma_c^2}. \qquad (5.9.64)$$

Acknowledgments

This research was supported in part by NSERC of Canada (Grant no. 8833) and Chinese Academy of Sciences President's International Fellowship Initiative (Grant no. 2016VBA039). We thank two anonymous referees for helpful comments to improve this paper.

Bibliography

[1] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[2] R. Axelrod, W.D. Hamilton, The evolution of cooperation. *Science* **211**, 1390-1396 (1981).

[3] C. Cannings, The latent roots of certain Markov chains arising in genetics: A new approach, I. Haploid models. *Advances in Applied Probabilioty* **6**, 260-290 (1974).

[4] W.J. Ewens, *Mathematical Population Genetics: I Theoretical Introduction* (Springer, New York, 2004).

[5] R.A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon, Oxford 1930).

[6] D. Fudenberg, D.G. Rand, A. Dreber, Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World. *American Economic Review* **102**(2), 720-749 (2012).

[7] C. Hilbe, Š. Šimsa, K. Chatterjee, M.A. Nowak, Evolution of cooperation in stochastic games. *Nature* **559**, 246-249 (2018).

[8] J. Hofbauer, K. Sigmund, *The Theory of Evolution and Dynamical Systems* (Cambridge University Press, Cambridge, 1998).

[9] H. Kaplan, K. Hill, A.M. Hurtado, “Risk, foraging and food sharing among the Ache” in *Risk and Uncertainty in Tribal and Peasant Economies*, pp. 107-144, E. Cashdan, Ed. (Westview Press, Boulder, CO, 1990).

[10] S. Karlin, B. Levikson, Temporal fluctuations in selection intensities: case of small population size. *Theoretical Population Biology* **6**, 383 (1974).

[11] M. Kimura, Diffusion models in population genetics. *Journal of Applied Probability* **1**(2), 177-232 (1964).

[12] R. Lande, S. Engen, B.-E. Saether, *Stochastic Population Dynamics in Ecology and Conservation* (Oxford University Press, Oxford, 2003).

[13] S. Lessard, Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory. *Theoretical Population Biology* **68**, 19-27 (2005).

[14] S. Lessard, "Evolution of cooperation in finite populations" in *Evolutionary Game Dynamics*, pp. 143-171, K. Sigmund, Ed. (American Mathematical Society, Providence, Rhode Island 2011).

[15] C. Li, T. Ji, Q-Q. He, X-D. Zheng, B-Y. Zhang, S. Lessard, Y. Tao, Uncertainty in the payoffs for defection could be conducive to the evolution of cooperation. *preprint* (2019).

[16] R.M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, 1973).

[17] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).

[18] M.A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).

[19] M.A. Nowak, R. Highfield, *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed* (Free Press, New York, 2011).

[20] M.A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646 (2004).

[21] M.A. Nowak, K. Sigmund, "How populations cohere: Five rules for cooperation" in *Theoretical Ecology - Principles and Applications*, pp. 7-16, R.M. May, A. McLean, Eds. (Oxford University Press, Oxford, 2007).

[22] H. Ohtsuki, C. Hauert, E. Lieberman, M.A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502 (2006).

[23] W. Poundstone, *Prisoner's Dilemma* (Doubleday, New York, 1992).

[24] H. Risken, *The Fokker-Planck Equation, Methods of Solution and Applications* (Springer, Berlin, 1992).

[25] E. Solan, N. Vieille, Stochastic games. *Proc. Natl. Acad. Sci. U.S.A.* **112**(45), 13743-13746 (2015).

[26] F. Stollmeier, J. Nagler, Unfair and anomalous evolutionary dynamics from fluctuating payoffs. *Physical Review Letters* **120**, 058101 (2018).

[27] P.D. Taylor, L. Jonker, Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* **40**, 145-156 (1978).

[28] R.L. Trivers, Evolution of reciprocal altruism. *Quarterly Review of Biology* **46**, 35-57 (1971).

[29] X-D. Zheng, C. Li, S. Lessard, Y. Tao, Evolutionary stability concepts in a stochastic environment. *Physical Review E* **96**, 032414 (2017).

[30] X-D. Zheng, C. Li, S. Lessard, Y. Tao, Environmental noise could promote stochastic local stability of behavioral diversity evolution. *Physical Review Letters* **120**, 218101 (2018).

[31] S. Wright, Evolution in Mendelian populations. *Genetics* **16**, 97-159 (1931).

Part II

Part 2: Games with opting-out

Chapter 6

Article 6

A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game

Xiu-Deng Zheng[1], Cong Li[2], Jie-Ru Yu[3], Shi-Chang Wang[1], Song-Jia Fan[1],
Bo-Yu Zhang[4] and Yi Tao[1]

[1]Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, P.R. China

[2]Department of Mathematics and Statistics, University of Montreal, Montreal, Canada

[3] College of Resources and Environmental Sciences, Gansu Agricultural University, Lanzhou, P.R. China

[4] School of Mathematical Science, Beijing Normal University, Beijing, P.R. China

Information

This paper appeared in *Journal of Theoretical Biology* **420** 12-17. It was received on 11 October 2016; revised on 23 February 2017; available online on 1 March 2017.

My contribution

Yi Tao and I initiated the study on long-term coexistence of cooperation and defection and established the basic model. Bo-Yu joined later on and suggested the method of fast equilibrium. I did the equilibrium analysis of the replicator dynamics, completed by results on finite populations by Xiu-Deng. I ran all the simulations.

Abstract

The long-term coexistence of cooperation and defection is a common phenomenon in nature and human society. However, none of the theoretical models based on the Prisoner's Dilemma (PD) game can provide a concise theoretical model to explain what leads to the stable coexistence of cooperation and defection in the long-term even though some rules for promoting cooperation have been summarized (Nowak, 2006, Science 314, 1560-1563). Here, based on the concept of direct reciprocity, we develop an elementary model to show why stable coexistence of cooperation and defection in the PD game is possible. The basic idea behind our theoretical model is that all players in a PD game prefer a cooperator as an opponent, and our results show that considering strategies allowing opting out against defection provide a general and concise way of understanding the fundamental importance of direct reciprocity in driving the evolution of cooperation.

Keywords: Evolution of cooperation; Long-term stable coexistence of cooperation and defection; Out-for-tat (OFT); Tit-for-tat (TFT).

6.1 Introduction

Five rules for promoting cooperation based on kin selection (Hamilton, 1964), direct and indirect reciprocity (Trivers, 1971; Axelrod and Hamilton, 1981; Axelrod, 1984; Nowak and Sigmund, 2005), graph selection (Nowak and May, 1992; Ohtsuki *et al.*, 2006) and group selection (Traulsen and Nowak, 2006a) have been summarized (Nowak, 2006a). The one-third law based on the stochastic evolutionary game in a finite population also shows how the emergence of cooperation can be favored by natural selection (Nowak, 2004). Although these theoretical results have been successful in explaining the evolution of cooperation, none of them provides a simple mechanism that can lead to stable coexistence of cooperation and defection in the long-term even though this phenomenon is common in nature and human society (Dugatkin, 1997).

Cooperation means that a donor pays a cost, c, for a recipient to get a benefit, b, where $b > c$ (Nowak, 2006b; Sigmund, 2010). In the corresponding one-shot Prisoner's Dilemma (PD) game, defection is the only Nash equilibrium (NE) (Nowak, 2006b; Sigmund, 2010). On the other hand, for the repeated PD game with two strategies TFT (tit-for-tat) and AllD (always defect), TFT is a NE if the expected number of iterated interactions between a pair of individuals is larger than the critical value $b/(b-c)$ (Axelrod and Hamilton, 1981; Axelrod, 1984; Nowak, 2006a, 2006b; Sigmund, 2010). However, the stable coexistence of TFT and AllD is impossible in the TFT-AllD game. Clearly, the success of TFT is mainly due to the increased chance of interactions between cooperators (Axelrod, 1984; Axelrod and Dion, 1988). That is, TFT provides a mechanism whereby cooperators preferentially interact among themselves. Similarly, assortative matching among cooperators has been used to explain why altruism can emerge (Eshel and Cavalli-Sforza, 1982; Cavalli-Sforza and Feldman, 1983; Fletcher and Doebeli, 2006; Taylor and Nowak, 2006; Pacheco *et al.*, 2008), although the evolutionary origin of the non-uniform interaction rates among cooperators has not been explained (Taylor and Nowak, 2006; Pacheco *et al.*, 2008). For the repeated PD game, one of the key assumptions is that the interaction between a pair of individuals will be repeated for several rounds, but that the expected number of iterated rounds is fixed (Axelrod, 1984; Axelrod and Dion, 1988; Nowak, 2006b; Sigmund, 2010). In particular, no player in a repeated PD game is able to unilaterally stop the interaction with his/her opponent. However, based on individual self-interest in the PD game, both cooperators and defectors prefer an opponent who cooperates (or only cooperators are always welcome). Thus, if players are able to unilaterally terminate the interactions with their opponents, then a simple rule will be followed by all individuals: I would like to keep my opponent if he/she is a cooperator; and if my opponent is a defector, I will immediately stop the interaction with him/her and seek a new partner instead. Clearly, this simple rule reflects the basic characteristics of direct reciprocity. Recently, an interesting study based on the concept of conditional dissociation, i.e. the option to leave an interacting partner in response to his/her behavior, found that a strategy called "out-for-tat" (OFT) may be important for the coexistence of cooperation and defection (Schuessler, 1989; Hayashi, 1993; Aktipis, 2004; Fujiwara-Greve and Okuno-Fujiwara, 2009; Izquierdo *et al.*, 2010, 2014). In this study, strategy OFT means that an individual will respond to defection by merely leaving, i.e. OFT will not tolerate defection but, unlike TFT, it does not seek revenge. Although this study shows a possibility for the coexistence of cooperation and defection because of OFT, it is still not clear what the dynamical mechanism of the coexistence is. To reveal the fundamental evolutionary force driving the coexistence of cooperation and defection, based only on the concept of direct reciprocity (Trivers, 1971; Axelrod and Hamilton, 1981; Axelrod, 1984), we develop a concise theoretical model to show how opting out against defection improves the coexistence of cooperation and defection in PD game settings.

6.2 Definitions and assumptions

Consider a simplified PD game with payoff matrix $\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}$ (Nowak, 2006b; Sigmund, 2010). Unlike the classic repeated game, we assume that the interaction between a pair of individuals can be continued but each player can unilaterally break off the interaction with his/her opponent at any time according to his/her own volition. This means that all individuals (including both cooperators and defectors) will respond to defection by merely leaving (i.e. all individuals use OFT) (Izquierdo *et al.*, 2010, 2014). On the other hand, we continue to assume as in the classic repeated game that the expected number of rounds between a pair of individuals is limited even if these two individuals would like to continue their interaction (Axelrod, 1984; Axelrod and Dion, 1988; Nowak, 2006b; Sigmund, 2010). Specifically, we assume that the interaction between a pair of individuals will be terminated after each round with probability ρ, where ρ is independent of these individuals' strategies. Thus, the probability that an interaction pair CC (where C represents cooperation) will remain in the next round is $1-\rho$, implying that the expected length of their interaction is $1/\rho$. On the other hand, the interaction pairs CD (where D denotes defection) and DD will never continue to the next round, becoming single individuals immediately. At the end of each round, all single individuals form new interaction pairs through random mating in the next round.

Let P_{CC}, P_{CD} and P_{DD} denote the frequencies of interaction pairs CC, CD and DD, respectively, with $P_{CC}+P_{CD}+P_{DD}=1$. Then, the frequency of C at time t, denoted by x, is given by $x = P_{CC}+P_{CD}/2$, and the frequency of D by $1-x = P_{CD}/2+P_{DD}$. Notice that, for a given population size N, the expected change of the frequency of cooperation from x to $x \pm 1/N$ in the time interval $\big(t, t+1/N\big)$ can be logically expressed as

$$\begin{aligned}
\langle \Delta x \rangle &\equiv \langle x(t+1/N) - x(t) \rangle \\
&= Pr\{\Delta x = 1/N\}\big(x(t) + 1/N\big) + Pr\{\Delta x = -1/N\}\big(x(t) - 1/N\big) \\
&\quad + \Big[1 - Pr\{\Delta x = 1/N\} - Pr\{\Delta x = -1/N\}\Big] x(t) - x(t) \\
&= \frac{1}{N}\Big[Pr\{\Delta x = 1/N\} - Pr\{\Delta x = -1/N\}\Big] , \qquad (6.2.1)
\end{aligned}$$

where $Pr\{\Delta x = \pm 1/N\}$ denotes the probability that Δx equals exactly $\pm 1/N$. On the other hand, notice also that the expected changes of numbers of interaction pairs CC , CD and DD are

$$N\left[(1-\rho)P_{CC} + \left(\frac{2\rho P_{CC} + P_{CD}}{2(\rho P_{CC} + P_{CD} + P_{DD})}\right)^2 \cdot \big(\rho P_{CC} + P_{CD} + P_{CC}\big) - P_{CC}\right] ,$$

$$N\left[\frac{(2\rho P_{CC} + P_{DD})(P_{CD} + P_{DD})}{2(\rho P_{CC} + P_{CD} + P_{DD})^2} \cdot \big(\rho P_{CC} + P_{CD} + P_{DD}\big) - P_{CD}\right] ,$$

$$N\left[\left(\frac{P_{CD} + 2P_{DD}}{2(\rho P_{CC} + P_{CD} + P_{DD})}\right)^2 \cdot \big(\rho P_{CC} + P_{CD} + P_{DD}\big) - P_{DD}\right] ,$$

respectively. Thus, the expected changes of P_{CC}, P_{CD} and P_{DD}, which are defined as $\Delta P_* = P_*(t+1/N) - P_*(t)$ for $* = CC$, CD and DD, are given by

$$\begin{aligned}
\langle \Delta P_{CC} \rangle &= (1-\rho)P_{CC} + \frac{(2\rho P_{CC} + P_{CD})^2}{4(1-(1-\rho)P_{CC})} - P_{CC} , \\
\langle \Delta P_{CD} \rangle &= \frac{(2\rho P_{CC} + P_{CD})(P_{CD} + 2P_{DD})}{2(1-(1-\rho)P_{CC})} - P_{CD} , \\
\langle \Delta P_{DD} \rangle &= \frac{(P_{CD} + 2P_{DD})^2}{4(1-(1-\rho)P_{CC})} - P_{DD} , \qquad (6.2.2)
\end{aligned}$$

respectively. Thus, for large N, the changes of P_{CC}, P_{CD} and P_{DD} should be considered to be the fast variables comparing to the change of x since $\lim_{N\to\infty} \langle \Delta x \rangle = 0$ but $\langle \Delta P_{CC} \rangle$, $\langle \Delta P_{CD} \rangle$ and $\langle \Delta P_{DD} \rangle$ are independent of N. Then, in analogy with the Hardy-Weinberg equilibrium in population genetics (Hofbauer and Sigmund, 1998), it is reasonable to assume that the interaction pairs CC, CD and DD are at a "temporal equilibrium" at any time t because of the random meeting between a pair of individuals. From the solutions of equations $\langle \Delta P_{CC} \rangle = 0$, $\langle \Delta P_{CD} \rangle = 0$ and $\langle \Delta P_{DD} \rangle = 0$, the temporal equilibrium satisfies $P_{CD}^2 = 4\rho P_{CC} P_{DD}$ (or $\big((1-\rho)/\rho\big) P_{CD}^2 + 2P_{CD} - 4x(1-x) = 0$ since $P_{CC} + P_{CD} + P_{DD} = 1$ and $x = P_{CC} + P_{CD}/2$). This implies that, at any time t, P_{CD} can be expressed as

$$P_{CD} = -\frac{\rho}{1-\rho} + \sqrt{\left(\frac{\rho}{1-\rho}\right)^2 + \frac{4x(1-x)\rho}{1-\rho}} \tag{6.2.3}$$

for all possible $0 < x < 1$ and $0 < \rho < 1$.

6.3 Stability analysis of the deterministic model

Based on the definitions and assumptions in Section 6.2, it is easy to see that, at any time t, a cooperator has an opponent displaying cooperation (respectively, defection) with probability $2P_{CC}\big/(2P_{CC} + P_{CD})$ $(P_{CD}\big/(P_{CD} + 2P_{CC})$, respectively). Similarly, a defector has an opponent displaying cooperation (respectively, defection) with probability $P_{CD}\big/(P_{CD} + 2P_{DD})$ $(2P_{DD}\big/(P_{CD} + 2P_{DD})$, respectively). This implies that the expected payoffs of C and D, denoted by π_C and π_D, respectively, can be expressed as

$$\begin{aligned} \pi_C &= \frac{2P_{CC}}{2P_{CC} + P_{CD}}(b-c) - \frac{P_{CD}}{2P_{CC} + P_{CD}} c \\ &= \frac{2x - P_{CD}}{2x} b - c \ , \\ \pi_D &= \frac{P_{CD}}{P_{CD} + 2P_{DD}} b \ = \ \frac{P_{CD}}{2(1-x)} b \ . \end{aligned} \tag{6.3.1}$$

Obviously, if the population size is assumed to be large enough, then the time evolution of x obeys a simple differential equation

$$\begin{aligned} \frac{dx}{dt} &= x(1-x)(\pi_C - \pi_D) \\ &= x(1-x)(b-c) - \frac{bP_{CD}}{2} \ , \end{aligned} \tag{6.3.2}$$

where P_{CD} is assumed to be at the temporal equilibrium (see Eq. (6.2.3)) (Hofbauer and Sigmund, 1998).

For the above differential equation, Eq. (6.3.2), it is easy to see that the boundary $x = 0$ must be at least locally asymptotically stable since $d(dx/dt)\big/dx|_{x=0} = -c$, and that the boundary $x = 1$ must be unstable since $d(dx/dt)\big/dx|_{x=1} = c$. On the other hand, it is also easy to see that the interior equilibrium of Eq. (6.3.2) is the solution of equation $\pi_C - \pi_D = 0$, i.e.

$$\begin{aligned} x(1-x)(b-c) - \frac{b}{2}\left[-\frac{\rho}{1-\rho} + \sqrt{\left(\frac{\rho}{1-\rho}\right)^2 + 4x(1-x)\frac{\rho}{1-\rho}}\,\right] &= 0 \\ \Rightarrow \quad x(1-x) &= \frac{bc}{(b-c)^2} \cdot \frac{\rho}{1-\rho} \ . \end{aligned} \qquad (6.3.3)$$

Thus, two interior equilibria (denoted by x_1^* and x_2^*, respectively, with $0 < x_2^* < x_1^* < 1$) exist if $\rho < (b-c)^2\big/(b+c)^2$, in which case

$$x_{1,2}^* \;=\; \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{bc}{(b-c)^2} \cdot \frac{\rho}{1-\rho}} \ ; \qquad (6.3.4)$$

$x_1^* = x_2^* = 1/2$ if $\rho = (b-c)^2\big/(b+c)^2$; and no interior equilibrium can exist if $\rho > (b-c)^2\big/(b+c)^2$. For $\rho = (b-c)^2\big/(b+c)^2$, the unique interior equilibrium $x^* = 1/2$ must be unstable since it is easy to see $dx/dt < 0$ for all possible $x \in (0,1)$ except for $x = 1/2$. On the other hand, for $\rho < (b-c)^2\big/(b+c)^2$, it is also easy to see that the interior equilibrium x_1^* (with $x_1^* > 1/2$) is locally asymptotically stable but x_2^* (with $x_2^* < 1/2$) is unstable since $dx/dt < 0$ for $x \in (x_1^*, 1)$, $dx/dt > 0$ for $x \in (x_2^*, x_1^*)$, and $dx/dt < 0$ for $x \in (0, x_2^*)$. All of these stability results not only show clearly a dynamical mechanism for the time evolution of cooperation when all individuals use OFT but also provide insights into how direct reciprocity leads to the coexistence of cooperation and defection.

Notice that in our model the pure strategy set is $\{C, D\}$, and that all individuals will respond to defection by using OFT. Thus, it is easy to see that the pure strategy D is an evolutionarily stable strategy (ESS) since, when all individuals use D, no mutant strategy can successfully invade this population; and it is also easy to see that the pure strategy C is not an ESS since, when all individuals use C, strategy D can successfully invade the population (Maynard Smith, 1982). On the other hand, when all individuals use a mixed strategy $(\varepsilon, 1-\varepsilon)$ (where an individual using this mixed strategy will display C with probability ε and D with probability $1-\varepsilon$), the proportions of interaction pairs CC, CD

and DD are ε^2, $2\varepsilon(1-\varepsilon)$ and $(1-\varepsilon)^2$, respectively, in each round (Broom and Rychtar, 2013); and the expected payoffs of C and D are $\pi_C = \varepsilon b - c$ and $\pi_D = \varepsilon b$, respectively (see Eq. (6.3.1)). Thus, the expected payoff of the mixed strategy $(\varepsilon, 1-\varepsilon)$ is $\varepsilon(b-c)$. This implies that no mixed strategy can be an ESS in our model. In particular, while the interior equilibrium x_1^* is locally asymptotically stable in the dynamics (6.3.2) when $\rho < (b-c)^2/(b+c)^2$, the mixed strategy $(x_1^*, 1-x_1^*)$ is not an ESS.

However, it may still be possible that an individual not using OFT successfully invades a population consisting of individuals using OFT. Notice that when a mutant not using OFT invades an OFT population, the probability that it will have an opponent using C, denoted by $\phi_{C|M}$, should be

$$\begin{aligned}\phi_{C|M} &= \frac{2\rho P_{CC} + P_{CD}}{2(\rho P_{CC} + P_{CD} + P_{DD})} \\ &= 1 - \frac{1-x}{\rho x + (1-x) + (1-\rho)P_{CD}/2}\end{aligned}$$

(i.e. this probability is the frequency of C in the group of single individuals at the end of each round). It is easy to see that we must have $\phi_{C|C} > \phi_{C|M}$ (or $\phi_{C|C} > x > \phi_{C|M}$). On the other hand, when the system state is at the stable interior equilibrium x_1^*, all OFT individuals have the same expected payoff. This means that the expected payoff of the mutant must be less than the expected payoff of any one OFT individual when $x = x_1^*$. Thus, when all individuals use OFT and the system state is at an interior stable equilibrium x_1^*, no individual not using OFT can successfully invade this population. This is illustrated by simulation results that show neither AllD nor TFT can invade a mixed population where they must play against both OFT-cooperators and OFT-defectors (Fig 6.1). The simulation is based on the standard Moran process (Zhou *et al.*, 2010; Broom and Rychtar, 2013), where the total population size is fixed with $N = 1000$. (In this process, in each time step only one individual is chosen for reproduction and one is chosen for elimination. In the population composed of AllD-individuals (or TFT-individuals), OFT-cooperators and OFT-defectors, let y_1, y_2 and y_3 be the frequencies of AllD-individuals (or TFT-individuals), OFT-cooperators and OFT-defectors, respectively, and π_1, π_2 and

π_3 denote the expected payoff of AllD-individual (TFT-individual), OFT-cooperator and OFT-defector, respectively. Then, in each time step, a new AllD-individual (or TFT-individual) is produced with probability $\pi_1 y_1 / \sum_{i=1}^{3} \pi_i y_i$, a new OFT-cooperator is produced with probability $\pi_2 y_2 / \sum_{i=1}^{3} \pi_i y_i$, and a new OFT-defector is produced with probability $\pi_3 y_3 / \sum_{i=1}^{3} \pi_i y_i$; and the probability that an AllD-individual (or TFT-individual), or an OFT-cooperator, or an OFT-defector, is chosen for elimination is y_1, or y_2, or y_3.) Since the population size is large (i.e. $N = 1000$), for any given initial state, as time passes the stochastic trajectories of OFT-C fluctuate around the stable equilibrium x_1^* obtained from Eq. (6.3.2).

6.4 Stochastic dynamics in a finite population

Notice that the boundary $x = 1$ (or $x = 0$) is always unstable (or stable) in dynamics (6.3.2). Thus, a natural question is whether the evolutionary emergence of cooperation will be favored by natural selection when all individuals use OFT. Specifically, for the situation with $\rho < (b-c)^2/(b+c)^2$ and the initial frequency of C in the interval $0 < x < x_2^*$, the question is whether the probability that the system state reaches (or passes) the stable equilibrium x_1^* at some time t is larger than the same probability under neutral selection. To show this, following previous works (Nowak *et al.*, 2004; Traulsen *et al.*, 2005; Nowak, 2006b), the stochastic dynamics is applied to our model in a finite population with fixed size N (where N could be large). We take the fitness of a cooperator (respectively, a defector) as $f_C = (1-\omega) + \omega\pi_C$ (respectively, $f_D = (1-\omega) + \omega\pi_D$), where the parameter ω denotes the selection intensity with $0 \leq \omega \leq 1$ (Nowak *et al.*, 2004; Nowak, 2006b). Following the approach of Traulsen *et al.* (2005), the transition probabilities that the system state changes from x to $x + 1/N$ and from x to $x - 1/N$ in a small time interval are defined as $s^+(x) = x(1-x)f_C/\bar{f}$ and $s^-(x) = x(1-x)f_D/\bar{f}$, respectively, where $\bar{f} = xf_C + (1-x)f_D$ denotes the mean fitness of the population.

Let $\phi(x; x_0, t)$ be the probability density distribution that the frequency of C equals x

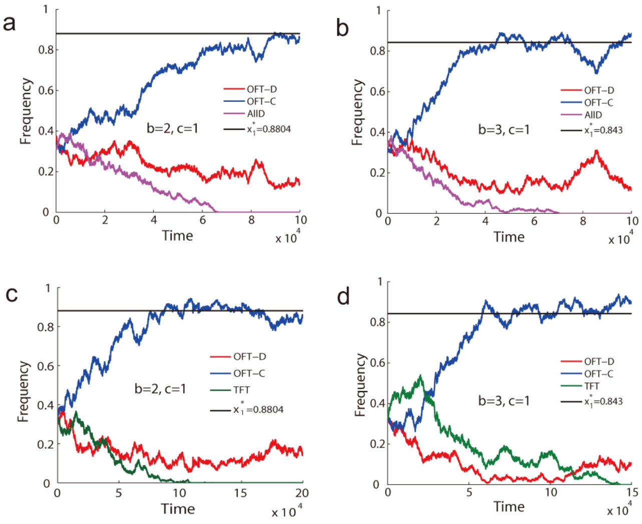


Figure 6.1: *Simulation of AllD (or TFT), OFT-cooperators and OFT-defectors in Moran Process.* The time evolution of AllD-individuals (or TFT-individuals), OFT-cooperators and OFT-defectors in a mixed population is simulated using the Moran Process (Nowak, 2006b), where the population size is fixed at $N = 1000$, the selection intensity is taken as $\omega = 0.01$ (Nowak, 2006b), the parameter ρ is taken as $\rho = 0.05$, and AllD-individuals (or TFT-individuals), OFT-cooperators and OFT-defectors have the same initial proportion, which is 1/3. In panels (**a**) and (**b**), the population consists of AllD-individuals, OFT-cooperators and OFT-defectors, where the payoff matrix is $\begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$ (with $x_1^* = 0.8804$) in panel (**a**), and $\begin{pmatrix} 2 & -1 \\ 3 & 0 \end{pmatrix}$ (with $x_1^* = 0.843$) in panel (**b**). Similarly, in panels (**c**) and (**d**), the population consists TFT-individuals, OFT-cooperators and OFT-defectors, where the payoff matrix in panel (**c**) is same as the payoff matrix in panel (**a**), and the payoff matrix in panel (**d**) is same as the payoff matrix in panel (**b**). All of these simulation results show clearly that neither AllD nor TFT can successfully invade a population composed of OFT-cooperators and OFT-defectors.

at time t when the initial frequency of C is x_0. Notice that

$$\begin{aligned}\phi(x;x_0,t+1/N)-\phi(x;x_0,t) &= \phi(x-1/N;x_0,t)s^+(x-1/N)\\ &+ \phi(x+1/N;x_0,t)s^-(x+1/N)\\ &- \phi(x;x_0,t)s^+(x)-\phi(x;x_0,t)s^-(x)\ ,\end{aligned} \tag{6.4.1}$$

and that the Taylor expansions of $\phi(x;x_0,t+1/N)$, $\phi(x\pm 1/N;x_0,t)$ and $s^\pm(x\mp 1/N)$ at x and t can be given by

$$\begin{aligned}\phi(x;x_0,t+1/N) &\approx \phi(x;x_0,t)+\frac{\partial}{\partial t}\phi(x;x_0,t)\frac{1}{N}\ ,\\ \phi(x\pm 1/N;x_0,t) &\approx \phi(x;x_0,t)\pm\frac{\partial}{\partial x}\phi(x;x_0,t)\frac{1}{N}+\frac{\partial^2}{\partial x^2}\phi(x;x_0,t)\frac{1}{2N^2}\ ,\\ s^\pm(x\mp 1/N) &\approx s^\pm(x)\mp\frac{\partial}{\partial x}s^\pm(x)\frac{1}{N}+\frac{\partial^2}{\partial x^2}s^\pm(x)\frac{1}{2N^2}\ .\end{aligned} \tag{6.4.2}$$

Thus, the Fokker-Planck equation (or diffusion approximation) of $\phi(x;x_0,t)$ can be expressed as

$$\frac{\partial\phi(x;x_0,t)}{\partial t} = \frac{\partial}{\partial x}D^{(1)}(x)\phi(x;x_0,t)+\frac{\partial^2}{\partial x^2}D^{(2)}(x)\phi(x;x_0,t)\ , \tag{6.4.3}$$

where $D^{(1)}(x)=s^+(x)-s^-(x)$ and $D^{(2)}(x)=\big(s^+(x)+s^-(x)\big)/2N$. Similarly, we have

$$\begin{aligned}\phi(x;x_0,t+1/N)-\phi(x;x_0,t) &= s^+(x_0)\phi(x;x_0+1/N,t)\\ &+ s^-(x_0)\phi(x;x_0-1/N,t)\\ &- s^+(x_0)\phi(x;x_0,t)-s^-(x_0)\phi(x;x_0,t)\ .\end{aligned} \tag{6.4.4}$$

Also from the Taylor expansions of $\phi(x;x_0,t+1/N)$ and $\phi(x;x_0\pm 1/N,t)$ at t and x_0, the backward Kolmogorov equation of $\phi(x;x_0,t)$ is given by

$$\frac{\partial\phi(x;x_0,t)}{\partial t} = D^{(1)}(x_0)\frac{\partial\phi(x;x_0,t)}{\partial x_0}+D^{(2)}(x_0)\frac{\partial^2\phi(x;x_0,t)}{\partial x_0^2}\ . \tag{6.4.5}$$

For an initial x_0 at time $t=0$, let $\mu(x_0,t)$ denote the probability that $x\geq x_1^*$ has occurred at or before time t, where $\mu(0,t)=0$ and $\mu(x_0,t)=1$ if $x_0\geq x_1^*$. From Eq. (6.4.5), the probability $\mu(x_0,t)$ also obeys the backward equation

$$\frac{\partial\mu(x_0,t)}{\partial t} = D^{(1)}(x_0)\frac{\partial\mu(x_0,t)}{\partial x_0}+D^{(2)}(x_0)\frac{\partial^2\mu(x_0,t)}{\partial x_0^2}\ . \tag{6.4.6}$$

The ultimate probability $\mu(x_0) = \lim_{t\to\infty} \mu(x_0, t)$ is the stationary solution of this equation, that is

$$D^{(1)}(x_0)\frac{d\mu(x_0,t)}{dx_0} + D^{(2)}\frac{d^2\mu(x_0,t)}{dx_0^2} \quad = \quad 0 \; .$$

Therefore, it can be expressed as

$$\mu(x_0) \quad = \quad \frac{\int_0^{x_0} \exp\left[-\int_0^z \frac{D^{(1)}(y)}{D^{(2)}(y)}dy\right] dz}{\int_0^{x_1^*} \exp\left[-\int_0^z \frac{D^{(1)}(y)}{D^{(2)}(y)}dy\right] dz} \tag{6.4.7}$$

(Traulsen *et al.*, 2006b; Zheng *et al.*, 2011). For the situation with $\omega N \ll 1$ (weak selection), the Taylor expansion of $\mu(x_0)$ about $\omega = 0$ is given by

$$\mu(x_0) \quad \approx \quad \frac{x_0}{x_1^*} + \omega N \frac{-x_1^* \int_0^{x_0} S(z)dz + x_0 \int_0^{x_1^*} S(z)dz}{(x_1^*)^2} \; , \tag{6.4.8}$$

where

$$S(z) \quad = \quad \int_0^z \left[(b-c) - \frac{b}{2} \cdot \frac{P_{CD}(y)}{y(1-y)}\right] dy \; . \tag{6.4.9}$$

So, it is easy to see that (i) $\mu(x_0) = x_0/x_1^*$ if $\omega = 0$ (i.e. under the neutral selection, $\mu(x_0)$ equals the ratio from x_0 to x_1^*); and (ii) for $\omega > 0$ but $\omega N \ll 1$, $\mu(x_0) > x_0/x_1^*$ if $x_0 \int_0^{x_1^*} S(z)dz > x_1^* \int_0^{x_0} S(z)dz$. In particular, for $x_0 = 1/N$, $\mu(1/N) > 1/(x_1^* N)$ if

$$\frac{1}{N}\int_0^{x_1^*} S(z)dz \quad > \quad x_1^* \int_0^{1/N} S(z)dz \; . \tag{6.4.10}$$

That is, under weak selection, the emergence of cooperation will be favored by natural selection if inequality (6.4.10) holds. The numerical analysis shows clearly that there exists a critical value of ρ, denoted by ρ_c, such that $\mu(1/N) > 1/x_1^* N$ if $\rho < \rho_c$ (see Fig 6.2). In this analysis, $\mu(x_0)$ is not defined as the fixation probability. Instead it is defined as the probability that the system will reach a state x_1^*, given that the initial state is x_0. Since x_1^* is locally asymptotically stable in the deterministic model, the condition $\mu(1/N) > 1/x_1^* N$ represents that the evolutionary emergence of cooperation will be favored by natural selection when $x_0 = 1/N$.

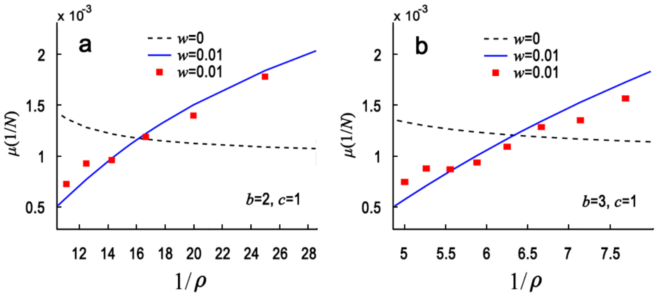


Figure 6.2: *The numerical analysis and stochastic simulation of* $\mu(1/N)$*.* The numerical analysis and stochastic simulation of $\mu(1/N)$, where the population size is fixed at $N = 1000$, the payoff matrix is $\begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$ in panel (**a**) and $\begin{pmatrix} 2 & -1 \\ 3 & 0 \end{pmatrix}$ in panel (**b**). In both panels (**a**) and (**b**), the dash curve corresponds to $\omega = 0$, blue curve denotes the numerical solution of $\mu(1/N)$ for $\omega = 0.01$, which is taken as the function of $1/\rho$, and the red squares are the results of stochastic simulation (using the Moran process (Nowak, 2006b)) for $\omega = 0.01$. These results show that there must exist a ρ_c such that $\mu(1/N) > 1/x_1^* N$ if $\rho < \rho_c$.

6.5 Discussion

In this paper, our analysis that emerges from allowing individuals to opt out against defection shows that this elementary mechanism based on direct reciprocity favors the emergence of cooperation by natural selection through the stochastic evolutionary game model and, at the same time, promotes the stable coexistence of cooperation and defection under standard evolutionary dynamics. This last result is especially important since stable coexistence is such a commonly observed phenomenon (Dugatkin, 1997) that does not occur for models of the repeated PD game that typically analyze such strategies as AllD, AllC in combination with others based on direct reciprocity (e.g. TFT, generous-TFT and win-stay lose-shift) (Nowak, 2006b, Sigmund, 2010). Besides, coexistence is possible even if the local game is a coordination game excluding the existence (Wu *et al.*, 2016). We agree with the generally recognized opinion that direct reciprocity is the most important force driving the evolution of cooperation, at least for the humans (Trivers, 1971; Clutton-Brock, 2002). Here, the strategies allowing opting out provide a better general mechanism to understand the fundamental importance of direct reciprocity.

We also noticed that for the long-term evolutionary dynamics of cooperation and defection, some previous studies (Nowak, 2006a, 2006b; Nowak and Sigmund, 2007) based on the repeated PD game suggested an evolutionary cycle of cooperation and defection (i.e. from AllD to TFT to GTFT to AllC and back to AllD) and that this cycle of cooperative and defective societies is a fundamental part of all observations regarding the evolution of cooperation. However, compared to these studies, our model not only provides a very simple mechanism such that the stable coexistence of cooperation and defection is possible but also an alternative possible explanation for the cooperative and defective behavior in nature and human society. Finally, we have to say that in our model the cost (or penalty) for changing partners is ignored. However, in real system the change of partners may have to pay some cost. The trade off between the cost for changing partners and the payoff may profoundly influence the evolution of cooperation. Thus, the effect of the cost for changing partners on the evolutionary dynamics of cooperation should be considered in future studies.

Bibliography

[1] W.D. Hamilton, The genetical evolution of social behaviour. *Journal of Theoretical Biology* **7**, 1-16 (1964).

[2] R. Trivers, The evolution of reciprocal altruism. *Quarterly Review of Biology* **46**, 35-57 (1971).

[3] R. Axelrod, W.D. Hamilton, The evolution of cooperation. *Science* **211**, 1390-1396 (1981).

[4] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[5] M.A. Nowak, K. Sigmund, Evolution of indirect reciprocity. *Nature* **437**, 1291-1298 (2005).

[6] M.A. Nowak, R.M. May, Evolutionary games and spatial chaos. *Nature* **359**, 826-829 (1992).

[7] H. Ohtsuki, C. Hauert, E. Lieberman, M.A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502-505 (2006).

[8] A. Traulsen, M.A. Nowak, Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 10952-10955 (2006a).

[9] M.A. Nowak, Five rules for the evolution of cooperation. *Science* **314**, 1560-1563 (2006a).

[10] M.A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646-650 (2004).

[11] L.A. Dugatkin, *Cooperation Among Animals: An Evolutionary Perspective* (Oxford University Press, New York, 1997).

[12] M.A. Nowak, Evolutionary Dynamics. (Belknap Press, Cambridge, 2006b).

[13] K. Sigmund, The Calculus of Selfishness. (Princeton University Press, Princeton, 2010).

[14] R. Axelrod, D. Dion, The further evolution of cooperation. *Science* **242**, 1385-1390 (1988).

[15] I. Eshel, L.L. Cavalli-Sforza, Assortment of encounters and evolution of cooperativeness. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 1331-1335 (1982).

[16] L.L. Cavalli-Sforza, M.W. Feldman, Paradox of the evolution of communication and of social interactivity. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 2017-2021 (1983).

[17] J.A. Fletcher, M. Doebeli, How altruism evolves: assortment and synergy. *Journal of Evolutionary Biology* **19**, 1389-1393 (2006).

[18] C. Taylor, M.A. Nowak, Evolutionary game dynamics with non-uniform interaction rates. *Theoretical Population Biology* **69**, 243-252 (2006).

[19] J.M. Pacheco, A. Traulsen, H. Ohtsuki, M.A. Nowak, Repeated games and direct reciprocity under active linking. *Journal of Theoretical Biology* **250**, 723-731 (2008).

[20] N. Hayashi, From TIT-for-TAT to OUT-for-TAT. *Sociol Theory Methods* **8**, 19-32 (1993).

[21] R. Schuessler, Exit threats and cooperation under anonymity. *Journal of Conflict Resolution* **33**, 728-749 (1989).

[22] C.A. Aktipis, Know when to walk away: contingent movement and the evolution of cooperation. *Journal of Theoretical Biology* **231**, 249-260 (2004).

[23] T. Fujiwara-Greve, M. Okuno-Fujiwara, Voluntarily separable repeated prisoner's dilemma. *The Review of Economic Studies* **76**, 993-1021 (2009).

[24] S.S. Izquierdo, L.R. Izquierdo, F. Vega-Redondo, The option to leave: Conditional dissociation in the evolution of cooperation. *Journal of Theoretical Biology* **267**, 76-84 (2010).

[25] S.S. Izquierdo, L.R. Izquierdo, F. Vega-Redondo, Leave and let leave: A sufficient condition to explain the evolutionary emergence of cooperation. *Journal of Economic Dynamics and Control* **46**, 91-113 (2014).

[26] J. Hofbauer, K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge Univ. Press, Cambridge, 1998).

[27] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).

[28] M. Broom, J. Rychtar, *Game-theoretical models in biology* (Crc Press, London UK, 2013).

[29] D. Zhou, B. Wu, H. Ge, Evolutionary stability and quasi-stationary strategy in stochastic evolutionary game dynamics. *Journal of Theoretical Biology* **264**, 874-881 (2010).

[30] A. Traulsen, J.C. Claussen, C. Hauert, Coevolutionary Dynamics: From finite to infinite populations. *Physical Review Letters* **95**, 238701 (2005).

[31] A. Traulsen, J.M. Pacheco, L.A. Imhof, Stochasitcity and evolutionary stability. *Physical Review E* **74**, 011901 (2006b).

[32] X.D. Zheng, R. Cressmn, Y. Tao, The diffusion approximation of stochastic evoltionary game dynamics: mean effective fixation time and the significance of the One-third Law. *Dynamic Games and Applications* **1**, 462-477 (2011).

[33] B. Wu, J. Arranz, J. Du, D. Zhou, A. Traulsen, Evolving synergetic interactions. *Journal of The Royal Society Interface* **13**, 20160282 (2016).

[34] T. Clutton-Brock, Breeding together: kin selection and mutualism in cooperative vertebrates. *Science* **296**, 69-72 (2002).

[35] M.A. Nowak, K. Sigmund, “How populations cohere: Five rules for cooperation” in *Theoretical Ecology - Principles and Applications*, pp. 7-16, R.M. May, A. McLean, Eds. (Oxford University Press, Oxford, 2007).

Chapter 7

Article 7

Opting out against defection leads to stable coexistence with cooperation

Bo-Yu Zhang[1,2,*], Song-Jia Fan[1,*], Cong Li[1,3,*], Xiu-Deng Zheng[1,*], Jian-Zhang Bao[4],
Ross Cressman[5] and Yi Tao[1]

[1]Key Lab of Animal Ecology and Conservation Biology, Chinese Academy of Science, Beijing, China.

[2]Laboratory of Mathematics and Complex Systems, Ministry of Education, School of Mathematical Sciences, Beijing Normal University, Beijing, China.

[3]Department of Mathematics and Statistics, University of Montreal, Montreal, Canada.

[4]School of Complex Systems, Beijing Normal University, Beijing, China.

[5]Department of Mathematics, Wilfrid Laurier University, Waterloo, Canada.

[*]These authors contributed equally to this work.

Information

This paper appeared in *Scientific Reports* **6** 35902. It was received on 4 March 2016; published on 24 October 2016.

My contribution

Bo-Yu and I did most of the theoretical analysis. Bo-Yu, Song-Jia and I designed the experiment with some advice from Yi Tao and performed the experiment.

Abstract

Cooperation coexisting with defection is a common phenomenon in nature and human society. Previous studies for promoting cooperation based on kin selection, direct and indirect reciprocity, graph selection and group selection have provided conditions that cooperators outcompete defectors. However, a simple mechanism of the long-term stable coexistence of cooperation and defection is still lacking. To reveal the effect of direct reciprocity on the coexistence of cooperation and defection, we conducted a simple experiment based on the Prisoner's Dilemma (PD) game, where the basic idea behind our experiment is that all players in a PD game should prefer a cooperator as an opponent. Our experimental and theoretical results show clearly that the strategies allowing opting out against defection are able to maintain this stable coexistence.

7.1 Introduction

A great deal of research has been devoted to explain how the evolution of cooperation can be favored by natural selection. Five rules for promoting cooperation based on kin selection (Hamilton, 1964), direct and indirect reciprocity (Trivers, 1971; Axelrod and Hamilton, 1981; Axelrod, 1984; Nowak and Sigmund, 2005), graph selection (Nowak and May, 1992; Ohtsuki *et al.*, 2006) and group selection (Traulsen *et al.*, 2006) have been summarized (Nowak, 2006a), and these models provided simple conditions that natural selection can lead to full cooperation. However, few literatures have considered how cooperation and defection can coexist in the long-term even though this phenomenon is common in nature and human society (Dugatkin, 1997). Other studies (Nowak, 2006b, Sigmund, 2010) have shown ongoing oscillations between cooperative and defective societies can evolve in theoretical models, possibly explaining such phenomena as the alternate appearance of war and peace (Nowak, 2006b). However, these models still do not provide a simple mechanism to drive the long-term stable coexistence of cooperation and defection.

Cooperation means that a donor pays a cost, c, for a recipient to get a benefit, b, where $b > c$ (Nowak, 2006b; Sigmund, 2010). In the corresponding one-shot PD game, defection is the only Nash equilibrium (NE) (Nowak, 2006b; Sigmund, 2010). On the other hand, for the repeated PD game with two strategies TFT (tit-for-tat) and AllD (always defect), TFT is a NE if the expected number of iterated interactions between a pair of individuals is larger than the critical value $b/(b-c)$ (Axelrod and Hamilton, 1981; Axelrod, 1984; Nowak, 2006a, 2006b; Sigmund, 2010). However, the stable coexistence of TFT and AllD is impossible in the TFT-AllD game. Clearly, the success of TFT is mainly due to the increased chance of interactions between cooperators (Axelrod, 1984; Axelrod and Dion, 1988). That is, TFT provides a mechanism whereby cooperators preferentially interact among themselves. Similarly, assortative matching among cooperators has been used to explain why altruism can emerge (Eshel and Cavalli-Sforza, 1982; Cavalli-Sforza and Feldman, 1983; Fletcher and Doebeli, 2006; Taylor and Nowak, 2006; Pacheco *et* *al.*, 2008), although the evolutionary origin of the non-uniform interaction rates among cooperators has not been explained (Taylor and Nowak, 2006; Pacheco *et al.*, 2008).

For the repeated PD game, one of the key assumptions is that the interaction between a pair of individuals will be repeated for several rounds, and no player in the game is able to stop the interaction with his/her opponent (Axelrod, 1984; Axelrod and Dion, 1988; Nowak, 2006b; Sigmund, 2010). However, based on individual self-interest in the PD game, both cooperators and defectors prefer an opponent who cooperates (i.e. only cooperator is always welcome). Thus, if players are able to unilaterally terminate the interactions with their opponents, then a simple rule will be followed by all individuals: I would like to keep my opponent if he/she is a cooperator; and if my opponent is a defector, I will stop the interaction with him/her and seek a new partner instead.

Recently, an interesting study based on the concept of conditional dissociation found that a strategy called out-for-tat (OFT) is important for the coexistence of cooperation and defection (Schuessler, 1989; Hayashi, 1993; Orbell and Dawes, 1993; Hauk, 2003; Aktipis, 2004; Fujiwara-Greve and Okuno-Fujiwara, 2009; Izquierdo *et al.*, 2010, 2014). Since OFT means that an individual displaying cooperation (C) will respond to defection (D) by merely leaving, OFT will not tolerate defection but, unlike TFT, it does not seek revenge. To reveal the fundamental evolutionary force driving the coexistence of C and D, we conduct a simple experiment based on the repeated PD game, where, unlike the classic repeated game, each player can unilaterally break off the pairwise interaction with his/her opponent according to his/her own volition. On the other hand, different from previous experiments on repeated PD game with outside option (Schuessler, 1989; Hayashi, 1993; Orbell and Dawes, 1993; Hauk, 2003; Aktipis, 2004; Fujiwara-Greve and Okuno-Fujiwara, 2009; Izquierdo *et al.*, 2010, 2014), the expected number of rounds between a pair of individuals is still limited in our experimental design even if these two individuals would like to continue their interaction (Axelrod, 1984; Axelrod and Dion, 1988; Nowak, 2006b; Sigmund, 2010).

7.2 Results

A total of 264 university students were divided into five groups, including two control groups (C1 and C2) and three treatment groups (T1, T2 and T3) (see **Appendix A1**). Note that the experimental settings in all three treatment groups T1, T2 and T3 are exactly the same, therefore in the data analysis we treat them as one group, denoted by T (see **Appendix A2**). The basic payoff matrix in our experiment is

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 4 & 1 \\ 5 & 2 \end{pmatrix} \end{array}, \tag{7.2.1}$$

where this payoff matrix can be normalized as a simplified PD game with $b = 3$ and $c = 1$. Each subject participated in 65 to 80 rounds of interactions between pairs of individuals playing this game over about 40 minutes. Participants were told that the experiment would be randomly stopped at 60-80 rounds. Thus, to avoid end-round effects and to keep the comparison unbiased, we only used data in the first 60 rounds in all groups in later statistical analysis.

The control experiments C1 and C2 are the classic repeated PD game and one-shot PD game, respectively. In C1, each interaction pair continues to the next round with probability 5/6 and is terminated with probability 1/6. At the end of each round, all single subjects form new interaction pairs through random meeting in the next round. In C2, all subjects are shuffled to form new interaction pairs in every round. On the other hand, the experimental setting in the treatment T is similar to C1 except that, at the end of each round, each subject decides whether he/she would like to continue the interaction in the next round with his/her current opponent. An interaction pair is terminated if at least one of the two subjects decides to stop; and is automatically terminated by the system with probability 1/6 even if both subjects choose to continue. After that, similarly, all single subjects are randomly repaired with a new opponent to play in the next round (see **Appendix A1**).

The primary experimental result (Fig. 7.1) is that for all control and treatment experiments, the cooperation level (defined as the frequency that C is used) in C1 is significantly higher than either in treatment T or in control C2, and that in T is significantly higher than in C2. (see **Appendix A2**).

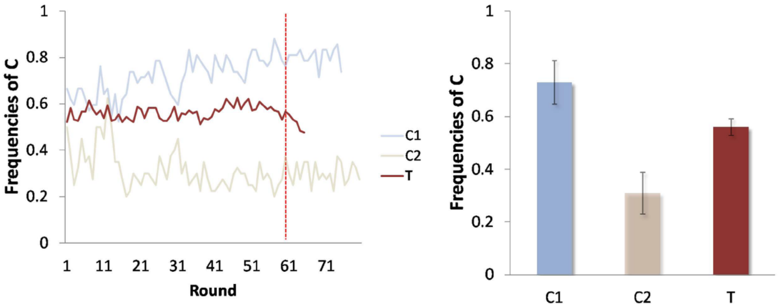


Figure 7.1: *Cooperation levels per round for treatment compared to control experiments.* Panel (a) shows the time evolution of cooperation levels per round in C1, C2 and T respectively, with dashed line at round 60. Panel (b) shows the average cooperation levels over 60 rounds with standard errors in C1, C2 and T, respectively, which are: 0.72 ± 0.0808 in C1; 0.32 ± 0.0876 in C2; and 0.56 ± 0.0287 in T. Mann-Whitney U-test shows that the differences between C1 and C2, between C1 and T and between T and C2 are significant with p-value < 0.01 (after Bonferroni correction) (see **Appendix A2**, Table 7.3).

It is also easy to see that the cooperation level in C1 increased over time from an average of 64% in the initial 10 rounds to an average of 80% from round 51 to 60. This time evolution of C in C1 can be characterized well by the TFT-AllD game, where both TFT and AllD are evolutionarily stable with current parameters (Nowak, 2006b, Sigmund, 2010). In particular, the system state tends to TFT given the initial data because it has a larger basin of attraction under the evolutionary dynamics. In C2, the cooperation level is much lower, declining from 39% in the initial 10 rounds to 28% from round 51 to 60. Obviously, C2 reflects well the characteristics of the one-shot PD game where only D is evolutionarily stable. However, in T, a relatively stable cooperation level is maintained over all rounds (e.g. an average of 56% in the initial 10 rounds and 58% from round 51 to 60). Moreover, the cooperation level in T is between the cooperation levels in C1 and in C2. This suggests that the treatment T provides a possible mechanism to maintain the stable coexistence of C and D.

In the treatment T, the chance that a subject decides to continue the interaction with his/her current opponent in the next round is 92% if his/her opponent displays C, whereas this chance drops to 53% if his/her opponent displays D (Fig. 7.2). We are also interested in how a player using strategy A responds when his/her opponent displays strategy B, where A, $B = C$, D. For the interaction pair C-C, only 10% of the interactions are stopped by the players, with 94% of the C-players choosing to continue the interaction with their current opponent; for the interaction pair C-D, the probability that at least one player chooses to stop the interaction is 56%, in which C-players (respectively, D-players) choose to stop the interaction with probability 33% (respectively 35%); and for the interaction pair D-D, 67% of the interactions are terminated by the players, in which each D-player chooses to stop the interaction with probability 43% (Fig. 7.2).

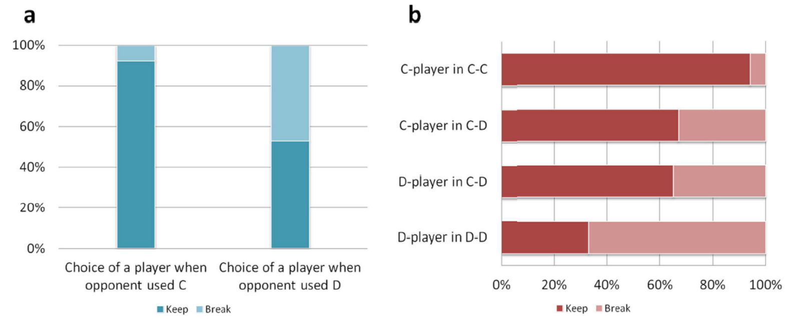


Figure 7.2: *Individuals' responses to the behavior of their opponents in the first 60 rounds.* Panel (a) shows the probability that, at the end of each round, a player chooses to keep, or break, the interaction with his/her opponent who uses C (D). Panel (b) shows the probabilities that, at the end of each round, a player using C (D) chooses to keep, or break, the interaction when his/her opponent uses C and when his/her opponent uses D.

We then identify the subjects who stop the interactions with significantly higher probability when their opponents display D than that when their opponents display C using two sided binomial sample test with 95% confidence intervals. These subjects are called OFT-strategists. According to this standard, most of the subjects (85.16%) can be classified as OFT-strategists. Only 3.3% of the subjects stop their interactions with significantly higher probability when their opponents display C than when they display D. The remaining 11.54% of the subjects cannot be identified (i.e., the chance that they will stop the interactions when their opponents display D is not significantly different from the chance when their opponents display C).

7.3 Discussion

To reveal the mechanism behind the treatment T maintaining the coexistence of C and D, we develop a concise theoretical framework to show how opting out against D leads to the coexistence of C in PD game settings. Consider a simplified repeated PD game (Nowak, 2006b, Sigmund, 2010) with payoff matrix $\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}$ (analysis for the general PD game is shown in **Appendix B5**). At the end of each round, each player can unilaterally break off the interaction with his/ her opponent according to his/her own volition. In fact, we assume that all individuals (including both cooperators and defectors) respond to D by merely leaving (i.e., using OFT) (Izquierdo *et al.*,2010,2014). Moreover, as in the classic repeated game, the interaction between a pair of individuals is terminated after each round with probability ρ even if these two individuals would like to continue their interaction (Axelrod, 1984; Axelrod and Dion, 1988; Nowak, 2006b; Sigmund, 2010). Thus, the probability that an interaction pair C-C continues in the next round is $1-\rho$, implying that the expected length of their interaction is $1/\rho$. On the other hand, the interaction pairs C-D and D-D will never continue to the next round, becoming single individuals immediately. At the end of each round, all single individuals form new interaction pairs through random meeting in the next round (Fig. 7.3).

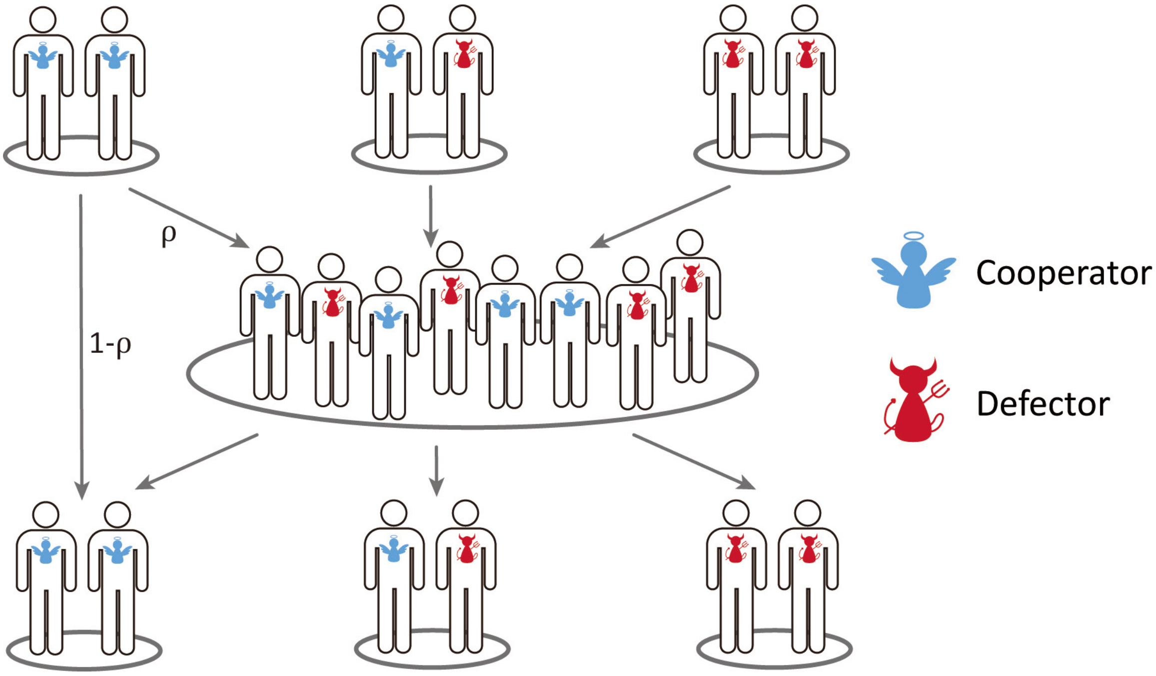


Figure 7.3: *The setup of the evolutionary model.* OFT-cooperators and OFT-defectors are marked by blue angels and red fiends, respectively. At the end of a round, *C*-*D* pairs and *D*-*D* pairs will be broken since all individuals immediately stop the interaction with a defector, and a *C*-*C* pair will be terminated with probability ρ even though both individuals are willing to continue. These single individuals will be paired with a new partner through random meeting in the next round.

Based on the theoretical analysis in **Appendix B4**, the time evolution of the frequency of OFT-cooperators, denoted by x, can be modeled by the replicator dynamics (Maynard Smith, 1982; Hofbauer and Sigmund, 1998)

$$\frac{dx}{dt} = x(1-x)(b-c) - \frac{bP_{CD}}{2}, \tag{7.3.1}$$

where the frequency P_{CD} of *C*-*D* pairs is shown to be given by Eq. (7.5.4) in **Appendix B4**. The stability analysis of this dynamics shows that (i) the boundary $x = 0$ is locally asymptotically stable for all possible $0 < \rho < 1$ but the boundary $x = 1$ is never stable; (ii) two interior equilibria (Eq. (7.5.6) in **Appendix B4**), denoted by x_1^* and x_2^* with $0 < x_2^* < 1/2 < x_1^* < 1$, exist if $\rho < (b-c)^2/(b+c)^2$, and x_1^* is locally asymptotically stable and x_2^*; (iii) a unique unstable interior equilibrium $x^* = 1/2$ exists if $\rho = (b-c)^2/(b+c)^2$; and (iv) the boundary $x = 0$ is globally asymptotically stable if $\rho > (b-c)^2/(b+c)^2$ (Fig. 7.4). Similar results for the general PD game are obtained in **Appendix B5**.

Clearly, there are some differences between this simple theoretical model and the experimental data in the treatments. In particular, the theoretical model assumes that all *C*-*D* pairs and *D*-*D* pairs will be terminated by the players, whereas in the experiments, the termination rate of such interaction pairs was 72% and subjects sometimes used *D* to response *D* (i.e., they adopted TFT-like strategies, see **Appendix A3** for details). Since more interactions with defectors were continued to the next round, it is therefore not surprising that the observed frequency of cooperation of 0.56 is less than the theoretically predicted stable equilibrium level x_1^* of 0.82 for our parameters (Fig. 7.4(a)). Nevertheless, the experiment and theory both show that adding the option of opting out can lead to the stable coexistence of *C* and *D*.

In conclusion, our experimental results and theoretical analysis that emerge from allowing individuals to opt out against defection show that this elementary mechanism based on direct reciprocity promotes the stable coexistence of cooperation and defection. These outcomes are especially important since stable coexistence is such a commonly observed phenomenon (Hamilton, 1964) that does not occur for models of the repeated PD game that typically analyze such strategies as AllD and AllC in combination with others based on direct reciprocity (e.g. TFT, generous-TFT and win-stay lose-shift) (Nowak, 2006b; Sigmund, 2010). This supports our contention that, while we agree with the generally recognized opinion that direct reciprocity is the most important force driving the evolution of cooperation (Axelrod, 1984; Clutton-Brock, 2002), strategies allowing opting out provide a better general framework for its analysis.

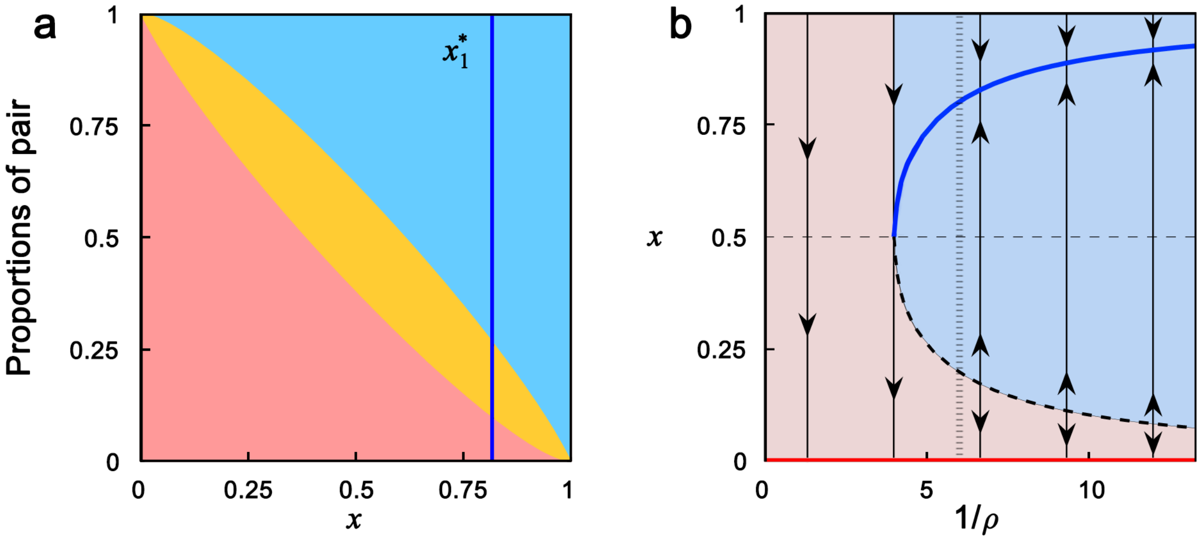


Figure 7.4: *Evolutionary dynamics of Eq. (7.3.1) with* $b = 3$ *and* $c = 1$*.* (a) Blue, yellow and pink regions represent respectively the proportions of C-C, C-D and D-D pairs for all possible $0 < x < 1$ at the temporal equilibrium with $P_{CD}^2 = 4\rho P_{CC} P_{DD}$ (Eq. (7.5.3) in **Appendix B4**), where the parameter ρ is taken as $\rho = 1/6$; and the blue line denotes the stable interior equilibrium $x_1^* = 0.82$. (b) Phase portrait of the dynamics Eq. (7.3.1) for different ρ. The red line denotes the stable boundary $x = 0$, the solid blue curve denotes the stable interior equilibrium x_1^* (which is bigger than $1/2$), and the dashed curve denotes the unstable interior equilibrium x_2^*. The population evolves to the boundary $x = 0$ for initial x in the pink region, and the dynamics leads to a stable coexistence of C and D for initial x in the blue region. The inverse $1/\rho$ represents the expected number of interactions of a C-C pair, where the vertical dash line denotes $1/\rho = 6$.

We now discuss some aspects of the experimental design and theoretical models, and review related literature.

In the experimental design, we assumed that the expected number of rounds in an interaction was limited even when there was an outside option because of the following two reasons. First, in the real world, an interaction may be terminated due to unexpected reasons even if both individuals would like to continue. In addition, this assumption allows us to compare the experimental results in T and C1 directly.

In the theoretical model, we assumed all individuals would respond to defection by leaving. However, many subjects in experiments used D as a response to D. A possible explanation of these behaviors would be direct reciprocity, e.g., these subjects want to punish their opponents by defection. Furthermore, there are also some subjects responded to D by C. They may expect that their kindness can encourage their opponent to cooperate in the future. Whatever the ultimate reasons behind these non-OFT behaviors are, it is also important to verify theoretically whether individuals not using OFT can successfully invade a population consisting of OFT-cooperators and OFT-defectors. It is easy to see that the expected payoff of an individual displaying C (or D) but not using OFT can be no higher than that of an OFT- cooperator (or OFT- defector) since the chance that an individual not using OFT is paired with an opponent displaying C will be less than that of an individual using OFT. Thus, when all individuals use OFT and the system state is at an interior stable equilibrium x_1^*, an individual not using OFT cannot successfully invade this population since OFT- cooperators and OFT- defectors have the same expected payoff at this equilibrium.

We note that there exist three classes of literature investigating the effect of outside option on cooperation, but their focus and results are different from ours. One class considers infinitely repeated PD game, where an interaction is terminated only if one of individuals in the partnership chooses to stop (Schuessler, 1989; Hayashi, 1993; Orbell and Dawes, 1993; Hauk, 2003; Aktipis, 2004; Fujiwara-Greve and Okuno-Fujiwara, 2009;

Vesely and Yang, 2012), or is deemed dead by the system (Izquierdo *et al.*, 2010, 2014). These studies often focus on Nash equilibrium (NE) or a certain class of strategies. For instance, it has been shown the game has no pure strategy NE because trust building strategies can defeat defectors (Fujiwara-Greve and Okuno-Fujiwara, 2009), and that TFT is dominated by some conditional and unconditional strategies (Izquierdo *et al.*, 2010, 2014). Another class allows abstaining from a game (Orbell and Dawes, 1993; Hauert *et al.*, 2002; Semmann *et al.*, 2003; Hauk, 2003;), with players choosing between an outside option (i.e., to be a loner) and the PD game. In such case, cooperators, defectors and loners can coexist if the payoff of the outside option is higher than the payoff of mutual defection. However, voluntary participation usually does not lead to a stable equilibrium, but to an unending limit cycle (Hauert *et al.*, 2002; Semmann *et al.*, 2003). The third class is developed on graph selection, arguing that dynamical networks where subjects can update their network connections can lead to cooperative outcomes (Santos *et al.*, 2006; Wu *et al.*, 2010; Fehl *et al.*, 2011; Rand *et al.*, 2011; Wang *et al.*, 2012; Bednarik *et al.*, 2014). When subjects play several PD games simultaneously with their neighbors, they often preferentially break social links with defectors and form new links with cooperators, which creates an incentive to cooperate (Fehl *et al.*, 2011; Rand *et al.*, 2011; Wang *et al.*, 2012; Bednarik *et al.*, 2014). Our model can be seen as the simplest dynamical networked PD game, where each individual only connects to one partner. However, network reciprocity, such as cooperators have more connections on average than defectors (Rand *et al.*, 2011; Wang *et al.*, 2012), or cooperators form large cooperative clusters (Fehl *et al.*, 2011; Bednarik *et al.*, 2014), are not included in our model.

Although the three classes of models contain the idea of walking away from the interaction with defectors, it cannot be simply concluded that outside option promotes the coexistence of cooperation and defection. Because other assumptions in these models, such as infinitely repeated game (Schuessler, 1989; Hayashi, 1993; Orbell and Dawes, 1993; Hauk, 2003; Aktipis, 2004; Fujiwara-Greve and Okuno-Fujiwara, 2009; Vesely and

Yang, 2012), optional participation (Hauert *et al.*, 2002; Semmann *et al.*, 2003) and spatial reciprocity (Santos *et al.*, 2006; Wu *et al.*, 2010; Fehl *et al.*, 2011; Rand *et al.*, 2011; Wang *et al.*, 2012; Bednarik *et al.*, 2014), may have positive effects on cooperation. Thus, our theoretical analysis and experimental results provide convincing evidences that opting out against defection alone is enough to maintain the stable coexistence of cooperation and defection.

There are also existing studies (Nowak, 2006b; Sigmund, 2010) that discuss the long-term evolutionary dynamics of cooperation and defection based on the repeated PD game. They have shown evolutionary cycles of cooperation and defection (i.e. from AllD to TFT to GTFT to AllC and back to AllD) can exist, and suggest that societal oscillation between cooperation and defection is a fundamental part of all our observations regarding the evolution of cooperation. However, compared to these studies, our model provides a very simple mechanism such that stable coexistence of cooperation and defection is possible without oscillation.

7.4 Methods: Experimental design

The experiments were conducted in computer labs at Beijing Normal University on April 2th and April 3th, 2015. The treatment group T1 was conducted on April 2th, and groups T2, T3, C1 and C2 were conducted on April 3th. All 264 participants were undergraduate students from Beijing Normal University who had no background in game theory and economics. The interactions between participants were anonymous, and via the computers. In the experiments, the participants were separated by the frosted-glass such that they could not see each other's computer screen, and they were not allowed to communicate during the experiment (see **Appendix A2**, Figure 7.5).

Before the experiment started, the rules of the game were explained to all participants, who were also shown the instructions of the experiment (in Chinese) for their particular control or treatment group. To ensure that all participants fully understand the game, they were required to answer correctly 4-5 questions before logging in to the formal experiments. The total number of rounds (or length) of the formal experiment for each of groups was about 65-80 (taking about 40 minutes), and the participants were told that the experiment would be randomly stopped at 60-80 rounds. Although there is no time limitation for participants' decision making in each round, it was recommended that participants submit their decisions within 30 seconds (there was a 30 second countdown on the screen).

When the experiment for each of groups was finished, the score of each participant in the experiment was converted to Chinese Yuan (CNY) with ratio 1: 0.3. The payoff of each participant plus a fixed amount of 20 Yuan was his/her total earning in the experiment. The overall average earning in our experiments was 83.9 Yuan (with minimum 63 Yuan, and maximum 108 Yuan); the group average earning is 80.6 Yuan (with minimum 63 Yuan, and maximum 95 Yuan) in T1, 79.3 Yuan (with minimum 66 Yuan, and maximum 91 Yuan) in T2, 83 Yuan (with minimum 72 Yuan, and maximum 96 Yuan) in T3, 98 Yuan (with minimum 83 Yuan, and maximum 108 Yuan) in C1, and 83 Yuan (with minimum 66 Yuan, and maximum 96 Yuan) in C2.

7.5 Appendix

A. Experimental Design and Results

7.5.1 Experimental design and description

The experiments based on the Prisoner's Dilemma (PD) game were conducted in 5 groups, of which 2 are control groups (denoted by C1 and C2 respectively), and 3 are treatment groups with the same experimental settings (denoted by T1, T2 and T3). The payoff matrix of the PD game is taken as

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 4 & 1 \\ 5 & 2 \end{pmatrix}, \end{array}$$

where C denotes the cooperation and D the defection. This payoff matrix can be normalized as

$$\begin{pmatrix} 4 & 1 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} 2 & -1 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix},$$

where $b = 3$ and $c = 1$. Therefore the benefit to cost ratio is $b/c = 3$.

Each subject in an interaction pair is to choose C or D in each round. At the end of each round, the players' choices and payoffs are shown to both of them (on the computer screen). For different groups, the special experimental designs are given below.

(1) C1 group follows the classic repeated PD games, in which each interaction pair is automatically stopped by the system at the end of each round with probability $\rho = 1/6$, and is continued in the next round with complementary probability $1 - \rho = 5/6$. If an interaction pair is automatically terminated, then both players are randomly re-paired with new opponents in the next round (Table 7.1). In the experiment, each round has at least 4 single subjects (note that each round has on average 7 single subjects), so we can guarantee that players will not meet their previous partner in the next round.

(2) C2 group follows the classic one-shot PD game, in which each interaction pair is terminated at the end of each round with probability $\rho = 1$. All players are then randomly re-paired in the next round (Table 7.1).

(3) Three treatment groups T1, T2 and T3 also follow the repeated PD game (similar to C1 group), but a new option is added for each player whereby he/she unilaterally stops the interaction with his/her opponent at the end of each round even if his/her opponent prefers to continue the game in next round. Thus, in three treatment groups (T1, T2 and T3), there are two reasons an interaction pair may stop at the end of each round. One is that the interaction is automatically stopped by the system with probability $\rho = 1/6$, and the other is that at least one player in the interaction pair unilaterally stops the interaction with his/her opponent. At the end of each round, not only the choice and payoff of each player and his/her opponent's choice and payoff are shown on the computer screen but also the reason why the interaction stops is shown when this occurs. If an interaction

	T1	T2	T3	C1	C2
Group size	60	66	56	42	40
Parameters	$b=3, c=1, \rho=1/6$			$b=3, c=1, \rho=1/6$	$b=3, c=1, \rho=1$
Option: Opting out	Yes			No	No
Number of rounds	65	65	65	75	80

Table 7.1: *Experimental design.*

pair is stopped, then both players are randomly re-paired with new opponents in the next round (Table 7.1). In the experiment, each round has at least 4 single subjects, so we can guarantee that players will not meet their previous partner in the next round.

7.5.2 Basic data analysis

The proportion of cooperation (C), denoted by P_C, in each round for each of C1, C2 and T (where T is the collection of T1, T2 and T3 since T1, T2 and T3 have the same experimental design) is shown in Figure 7.5a (i.e. time evolution of C) (see also Figure 7.1 in the main text), where the average of P_C in the first 60 rounds is 0.72 in C1, 0.32 in C2 and 0.56 in T (Table 7.2, Figure 7.6).For the average of P_C, C1 is significantly larger than T, and T is significantly larger than C2 (Table 7.3). From Figure 7.5a, it is clear that the cooperation level (P_C) slowly increases from 0.64 to 0.8 in C1, slowly decreases from 0.39 to 0.28 in C2, and maintains a constant level of about 0.56 in T. Furthermore, to show the cooperation level in each of T1, T2 and T3, the proportion of C per round for each of T1, T2 and T3 is also shown in Figure 7.7a, where the average of P_C in the first 60 rounds is 0.56 in T1, 0.52 in T2 and 0.62 in T3 (Table 7.2), and the differences between T1 and T2, between T1 and T3, and between T2 and T3 are not significant (Table 7.3, Figure 7.6).

Notice that, for each interaction pair, there are three possible strategy-pairs in each round, which are C-C, C-D and D-D, respectively (for example, if one player displays C and his/her opponent also displays C, then this strategy-pair is denoted by C-C). For convenience, the proportions of interaction pairs C-C, C-D and D-D in each round

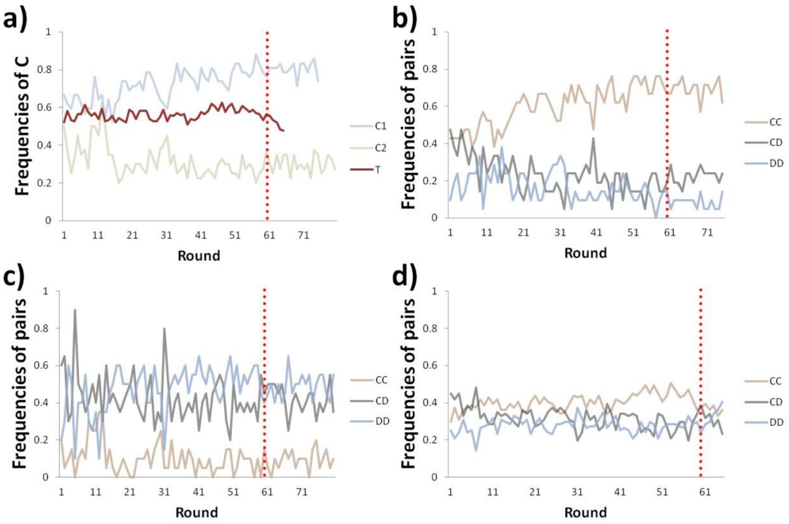


Figure 7.5: *P_C and the frequency of pairs in C1, C2 and T.* Panel (a) shows P_C per round for each of C1, C2 and T. Panels (b), (c) and (d) show P_{CC}, P_{CD} and P_{DD} per round in C1, C2 and T, respectively. The dotted lines mark at round 60.

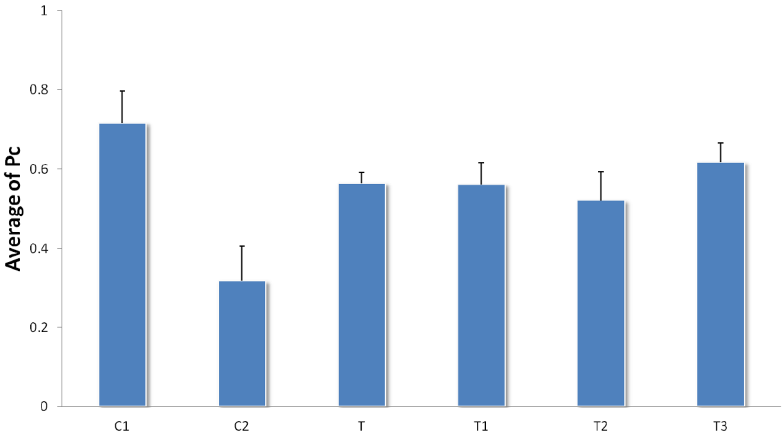


Figure 7.6: *Average P_C in C1, C2 and T (T1, T2, T3).* Average P_C in C1, C2 and T (T1, T2, T3) with standard errors in the first 60 rounds, which is 0.72 in C1, 0.32 in C2, 0.56 in T (0.56 in T1, 0.52 in T2, 0.62in T3).

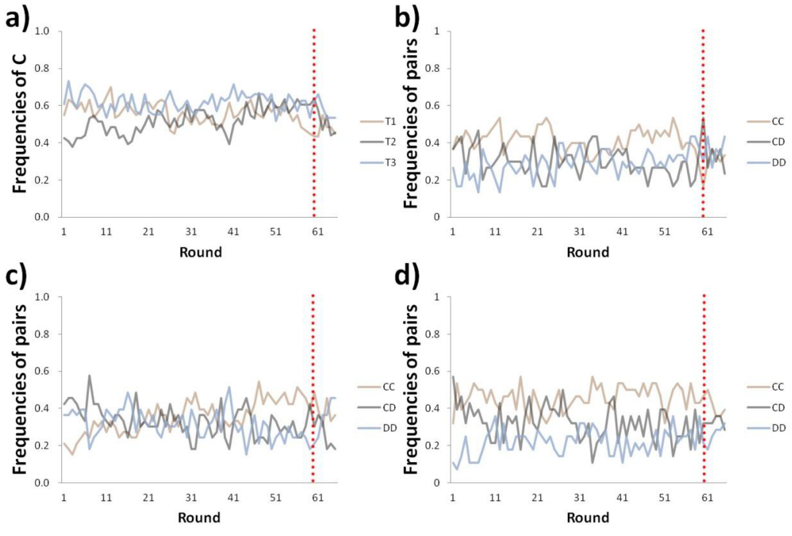
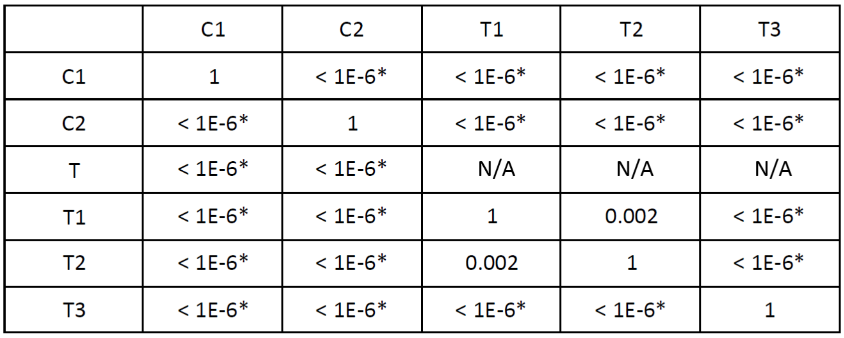


Figure 7.7: *P_C and the frequency of pairs in T1, T2 and T3.* Panel (a) shows P_C per round for each of T1, T2 and T3. Panels (b), (c) and (d) show P_{CC}, P_{CD} and P_{DD} per round in T1, T2 and T3, respectively. The dotted lines mark at round 60.

Average	C1	C2	T	T1	T2	T3
P_C	0.72	0.32	0.56	0.56	0.52	0.62
P_{CC}	0.60	0.10	0.40	0.41	0.35	0.46
P_{CD}	0.23	0.43	0.32	0.30	0.33	0.32
P_{DD}	0.17	0.47	0.28	0.29	0.31	0.22

Table 7.2: *Average of P_C, and averages of P_{CC}, P_{CD} and P_{DD} for each of C1, C2, T (T1, T2, T3).*



	C1	C2	T1	T2	T3
C1	1	< 1E-6*	< 1E-6*	< 1E-6*	< 1E-6*
C2	< 1E-6*	1	< 1E-6*	< 1E-6*	< 1E-6*
T	< 1E-6*	< 1E-6*	N/A	N/A	N/A
T1	< 1E-6*	< 1E-6*	1	0.002	< 1E-6*
T2	< 1E-6*	< 1E-6*	0.002	1	< 1E-6*
T3	< 1E-6*	< 1E-6*	< 1E-6*	< 1E-6*	1

Table 7.3: *Statistical test for difference in the average of P_C between each group.* Mann-Whitney U-test for difference in the average of P_C ($n_1 = 60$, $n_2 = 60$) between C1 and C2, between C1 and T, between C2 and T, between T1 and T2, between T1 and T3, and between T2 and T3, where the symbol "*" denotes that the difference is significant at $\alpha = 0.01/12 = 8.3E - 4$ (with Bonferroni correction).

are denoted by P_{CC}, P_{CD} and P_{DD}, respectively. Similar to 1 the analysis of P_C, the proportions of interaction pairs C-C, C-D and D-D (i.e. P_{CC}, P_{CD} and P_{DD}) per round for each of C1, C2 and T are plotted in Figure 7.5b, in Figure 7.5c and in Figure 7.5d, respectively. The averages of P_{CC}, P_{CD} and P_{DD} in the first 60 rounds are 0.60, 0.23 and 0.17 in C1; 0.10, 0.43 and 0.47 in C2; and 0.40, 0.32 and 0.28 in T (Table 7.2, Figure 7.8). For the averages of P_{CC}, P_{CD} and P_{DD} in each of C1, C2 and T, the difference between C1 and T and the difference between C2 and T are all significant (Table 7.4). Furthermore, for each of T1, T2 and T3, the proportions P_{CC}, P_{CD} and P_{DD} per round are plotted in Figure 7.7b, Figure 7.7c and Figure 7.7d, respectively. The averages of P_{CC}, P_{CD} and P_{DD} are 0.41, 0.30 and 0.29 in T1; 0.35, 0.33 and 0.31 in T2; and 0.46, 0.32 and 0.22 in T3 (Table 7.2, Figure 7.8, Table 7.4).

(a) Mann-Whitney U-test for difference in the average of P_{CC}

	C1	C2	T1	T2	T3
C1	1	< 1E-6*	< 1E-6*	< 1E-6*	< 1E-6*
C2	< 1E-6*	1	< 1E-6*	< 1E-6*	< 1E-6*
T	< 1E-6*	< 1E-6*			
T1	< 1E-6*	< 1E-6*	1	9.6E-5*	0.004
T2	< 1E-6*	< 1E-6*	9.6E-5*	1	< 1E-6*
T3	< 1E-6*	< 1E-6*	0.004	< 1E-6*	1

(b) Mann-Whitney U-test for difference in the average of P_{CD}

	C1	C2	T1	T2	T3
C1	1	< 1E-6*	3.5E-5*	< 1E-6*	1E-6*
C2	< 1E-6*	1	< 1E-6*	2.5E-5*	6E-6*
T	< 1E-6*	< 1E-6*			
T1	3.5E-5*	< 1E-6*	1	0.015	0.199
T2	< 1E-6*	2.5E-5*	0.015	1	0.614
T3	1E-6*	6E-6*	0.199	0.614	1

(c) Mann-Whitney U-test for difference in the average of P_{DD}

	C1	C2	T1	T2	T3
C1	1	< 1E-6*	< 1E-6*	< 1E-6*	1.1E-4*
C2	< 1E-6*	1	< 1E-6*	< 1E-6*	< 1E-6*
T	< 1E-6*	< 1E-6*			
T1	< 1E-6*	< 1E-6*	1	0.061	1E-6*
T2	< 1E-6*	< 1E-6*	0.061	1	< 1E-6*
T3	1.1E-4*	< 1E-6*	1E-6*	< 1E-6*	1

Table 7.4: *Statistical test for difference in the averages of* P_{CC}*,* P_{CD} *and* P_{DD} *between each group.* Mann-Whitney U-test for difference in the averages of P_{CC}, P_{CD} and P_{DD} ($n_1 = 60$, $n_2 = 60$) between C1 and C2, between C1 and T, between C2 and T, between T1 and T2, between T1 and T3, and between T2 and T3, where the symbol "*" denotes that the difference is significant at $\alpha = 0.01/12 = 8.3E - 4$ (with Bonferroni correction).

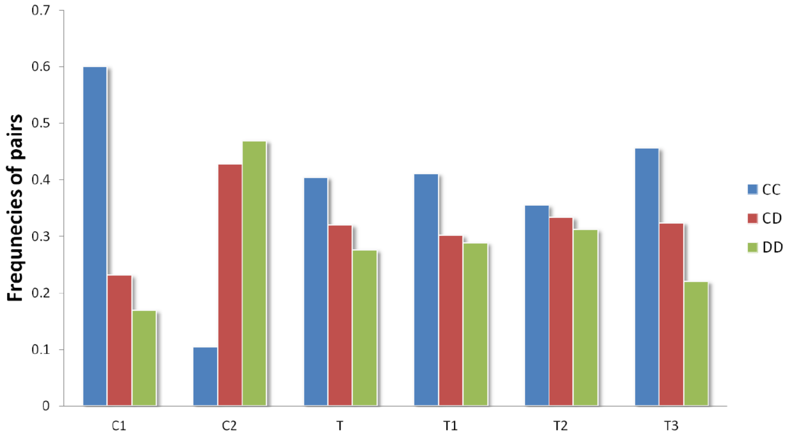


Figure 7.8: *Averages of* P_{CC}*,* P_{CD} *and* P_{DD} *in C1, C2 and T (T1, T2, T3).* Averages of P_{CC}, P_{CD} and P_{DD} in C1, C2 and T (T1, T2, T3) in the first 60 rounds, which are 0.60, 0.23, 0.17 in C1; 0.10, 0.43, 0.47 in C2; 0.40, 0.32, 0.28 in T (0.41, 0.30, 0.29 in T1; 0.35, 0.33, 0.31 in T2; 0.46, 0.32, 0.22 in T3).

7.5.3 Individual's response to his/her opponent's behavior in treatment T

To show how a player responds to his/her opponent's behavior (C or D), individual responses (i.e. continue the interaction with the current opponent in the next round, or stop the interaction with the current opponent) to C and D in T1, in T2 and in T3 are shown in Table 7.5a. Here, the response to C (D) is measured by the probability that the interaction will be kept, or will be stopped. It is easy to see that individuals' responses to C are very similar between T1, T2 and T3 (Table 7.6). The responses to D are also similar between T1 and T2, but are different between T1 and T3, and between T2 and T3 (Table 7.6).

Furthermore, we look into how a player using strategy A responds if his/her opponent displays strategy B, where A, $B = C$, D. The probabilities that a player using strategy A chooses to keep, or stop, the interaction with his/her opponent using strategy B (where

A, $B = C$, D) in the treatments (T1, T2, and T3) are shown in Table 7.5b. It is easy to see that in all the treatment groups, the probability of a player to choose keep is much higher if the opponent displays C rather than D. In Table 7.5c we show the probability that players in a particular interaction pair choose to stop or keep, and if the interaction is continued, what the choice combination in the next round is going to be. For the C-C pair, only 10% of the interactions are terminated by choice, and of all the pairs continue to the next round, 88% display C-C, 10% C-D, and only 2% D-D. For C-D pairs, the interactions are stopped by choice at a probability of 56%; and in the next round, 25% of the continued pairs display C-C, 23% C-D, 19% D-C, and 33% D-D. The probability that a D-D interaction pair is stopped by choice is 67%; in the next round, 61% of the interaction pairs stay at D-D, only 6% turn into C-C, and 32% to C-D.

B. Theoretical Analysis

7.5.4 Evolutionary dynamics for the PD game with additive payoff matrix , where all individuals use OFT

For payoff matrix $\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}$, let P_{CC}, P_{CD} and P_{DD} denote the proportions of interaction pairs C-C, C-D and D-D, respectively. Then, the frequency of C, denoted by x, is given by $x = P_{CC} + P_{CD}/2$, and the frequency of D is $1 - x = P_{CD}/2 + P_{DD}$. For a large population, the changes in proportions P_{CC}, P_{CD} and P_{DD} between rounds should be considered as fast variables comparing to the change of x since the meeting between a pair of individuals is random. To show this, consider a large population with size N, in which the expected change from x to $x \pm 1/N$ in the time interval $(t, t + 1/N)$, denoted

(7.5a) The probabilities that a player chooses to keep, or break, the interaction with his/her opponent using strategy A at the end of each round, where $A = C$, D.

Group	Choice of a player when his\her opponent uses C		Choice of a player when his\her opponent uses D	
	Keep	Break	Keep	Break
T1	0.91	0.09	0.56	0.44
T2	0.92	0.08	0.56	0.44
T3	0.92	0.08	0.45	0.55
T	0.92	0.08	0.53	0.47

(7.5b) The probabilities that a player using strategy A chooses to keep, or break, the interaction with his/her opponent using strategy B at the end of each round, where A, B = C, D.

Group	Choice of a player using C in C-C		Choice of a player using C in C-D		Choice of a player using D in D-C		Choice of a player using D in D-D	
	Keep	Break	Keep	Break	Keep	Break	Keep	Break
T1	0.96	0.04	0.76	0.24	0.59	0.41	0.55	0.45
T2	0.94	0.06	0.62	0.38	0.72	0.28	0.60	0.40
T3	0.92	0.08	0.62	0.38	0.62	0.38	0.54	0.46
T	0.94	0.06	0.67	0.33	0.65	0.35	0.57	0.43

(7.5c) The probability that an interaction pair C-C (or C-D, D-D) is broken at the end of each round, and the probability that two players display a particular strategy-pair in the next round if the interaction between these two players is kept.

<table>
<tr><th rowspan="2">Interaction pair</th><th rowspan="2">Probability that an interaction pair is stopped</th><th colspan="4">Probabilities that two subjects display interaction pairs C-C, C-D and D-D in the next round if the interaction between these two subjects is kept</th></tr>
<tr><th>C-C</th><th>C-D</th><th>D-C</th><th>D-D</th></tr>
<tr><td>C-C</td><td>0.13</td><td>0.88</td><td colspan="2">0.05</td><td>0.02</td></tr>
<tr><td>C-D</td><td>0.68</td><td>0.25</td><td>0.23</td><td>0.19</td><td>0.33</td></tr>
<tr><td>D-D</td><td>0.77</td><td>0.06</td><td colspan="2">0.16</td><td>0.61</td></tr>
</table>

Table 7.5: *The probabilities for breaking pairs in the experiment.*

	Test for difference in the average frequencies that a player chooses to keep the interaction after his/her opponent uses C			Test for difference in the average frequencies that a player chooses to keep the interaction after his/her opponent uses D		
	T1	T2	T3	T1	T2	T3
T1	1	0.197	0.262	1	0.902	< 0.001*
T2	0.197	1	0.698	0.902	1	< 0.001*
T3	0.262	0.698	1	< 0.001*	< 0.001*	1

Table 7.6: *Statistical test in the average frequencies that a player chooses to keep the interaction.* Mann-Whitney U-test for difference in the average frequencies that a player chooses to keep the interaction after his/her opponent uses strategy A in that round ($n_1 = 60$, $n_2 = 60$) between T1 and T2, between T2 and T3, and between T1 and T3, where A=(C, D). The symbol "*" denotes that the difference is significant at $\alpha = 0.01/3 = 3.3E - 3$ (with Bonferroni correction).

by $\Delta x = x(t + 1/N) - x(t)$, is

$$\begin{aligned}
\langle \Delta x \rangle &= \langle x(t + 1/N) - x(t) \rangle \\
&= Pr(\Delta x = 1/N)(x(t) + 1/N) + Pr(\Delta x = -1/N)(x(t) - 1/N) \\
&\quad + [1 - Pr(\Delta x = 1/N) - Pr(\Delta x = -1/N)]x(t) - x(t) \\
&= \frac{1}{N}\left[Pr(\Delta x = 1/N) - Pr(\Delta x = -1/N)\right]. \qquad (7.5.1)
\end{aligned}$$

Where $Pr(\Delta x = \pm 1/N)$ denotes the probability that Δx equals exactly $\pm 1/N$. Similarly, the expected changes to P_{CC}, P_{CD} and P_{DD} are given by

$$\begin{aligned}
\langle \Delta P_{CC} \rangle &\approx (1-\rho)P_{CC} + \frac{(2\rho P_{CC} + P_{CD})^2}{4(1-(1-\rho)P_{CC})} - P_{CC}, \\
\langle \Delta P_{CD} \rangle &\approx \frac{(2\rho P_{CC} + P_{CD})(P_{CD} + 2P_{DD})}{2(1-(1-\rho)P_{CC})} - P_{CD}, \\
\langle \Delta P_{DD} \rangle &\approx \frac{(P_{CD} + 2P_{DD})^2}{4(1-(1-\rho)P_{CC})} - P_{DD}. \qquad (7.5.2)
\end{aligned}$$

respectively. Notice that $\langle \Delta P_{CC} \rangle$, $\langle \Delta P_{CD} \rangle$ and $\langle \Delta P_{DD} \rangle$ are independent of population size N, and that $\lim_{N\to\infty} \langle \Delta x \rangle = 0$. Thus, for large N, the changes of P_{CC}, P_{CD} and P_{DD}

are fast variables comparing to the change of x. Thus, in analogy to the Hardy-Weinberg equilibrium in population genetics (see Hofbauer and Sigmund, 1998), we can assume that at any time t the proportions P_{CC}, P_{CD} and P_{DD} are at a "temporal equilibrium" for the current value of x, satisfying

$$P_{CD}^2 = 4\rho P_{CC} P_{DD}. \tag{7.5.3}$$

Therefore, we obtain

$$P_{CD} = -\frac{\rho}{1-\rho} + \sqrt{\left(\frac{\rho}{1-\rho}\right)^2 + \frac{4x(1-x)\rho}{1-\rho}}. \tag{7.5.4}$$

for all $0 < x < 1$ and $0 < \rho < 1$ since $P_{CC} + P_{CD} + P_{DD} = 1$ and $x = P_{CC} + P_{CD}/2$ (see Figure 7.4 in the main text).

Notice that, at any time t, an individual using C has an opponent displaying C (respectively, D) with probability $2P_{CC}/(2P_{CC} + P_{CD})$ (respectively, $P_{CD}/(2P_{CC} + P_{CD})$). Similarly, an individual using D has an opponent displaying C (respectively, D) with probability $P_{CD}/(P_{CD} + 2P_{DD})$ (respectively, $2P_{DD}/(P_{CD} + 2P_{DD})$). The expected payoffs of C and D, denoted by π_C and π_D, respectively, are then given by

$$\begin{aligned} \pi_C &= \frac{2P_{CC}}{2P_{CC} + P_{CD}}(b-c) - \frac{P_{CD}}{2P_{CC} + P_{CD}}c = \frac{2x - P_{CD}}{2x}b - c, \\ \pi_D &= \frac{P_{CD}}{P_{CD} + 2P_{DD}}b = \frac{P_{CD}}{2(1-x)}b. \end{aligned} \tag{7.5.5}$$

Thus, the time evolution of x can be given by $dx/dt = x(1-x)(\pi_C - \pi_D)$, i.e., Eq. (7.3.1) in the main text.

The boundary $x = 0$ of Eq. (7.3.1) in the main text is at least locally asymptotically stable since $d(dx/dt)/dt|_{x=0} = -c$, but the boundary $x = 1$ must be unstable since $d(dx/dt)/dt|_{x=1} = c$. On the other hand, an interior equilibrium of Eq. (7.3.1) in the main text must satisfy $\pi_C - \pi_D = 0$, i.e., $x(1-x) = (bc/(b-c)^2)(\rho/(1-\rho))$. Thus, two interior equilibria, denoted by x_1^* and x_2^*, respectively, with $0 < x_2^* < x_1^* < 1$, exist if $\rho < (b-c)^2/(b+c)^2$, in which case

$$x_{1,2}^* = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{bc\rho}{(b-c)^2(1-\rho)}}. \tag{7.5.6}$$

$x_1^* = x_2^* = 1/2$ if $\rho = (b-c)^2/(b+c)^2$; and no interior equilibrium can exist if $\rho > (b-c)^2/(b+c)^2$. For $\rho = (b-c)^2/(b+c)^2$, the unique interior equilibrium $x^* = 1/2$ must be unstable since $dx/dt < 0$ for all $x \in (0,1)$ except for $x = 1/2$. On the other hand, for $\rho < (b-c)^2/(b+c)^2$, the interior equilibrium $x_1^* > 1/2$ is locally asymptotically stable but x_2^* is unstable since $d(dx/dt)/dt|_{x=x_1^*} < 0$ and $d(dx/dt)/dt|_{x=x_2^*} > 0$.

7.5.5 Analysis of PD game with general payoff matrix, where all players use OFT

For the payoff matrix $\begin{pmatrix} R & S \\ T & S \end{pmatrix}$ with $T > R > P > S$, similar to the analysis in the subsection above, the expected payoffs π_C and π_D are given by

$$\begin{aligned} \pi_C &= \frac{2P_{CC}}{2P_{CC}+P_{CD}}R + \frac{P_{CD}}{2P_{CC}+P_{CD}}S = \frac{2x-P_{CD}}{2x}R + \frac{P_{CD}}{2x}S, \\ \pi_D &= \frac{P_{CD}}{P_{CD}+2P_{DD}}T + \frac{2P_{DD}}{P_{CD}+2P_{DD}}P = \frac{P_{CD}}{2(1-x)}T + \frac{2(1-x)-P_{CD}}{2(1-x)}P. \end{aligned} \tag{7.5.7}$$

Thus, the time evolution of x can be described by

$$\frac{dx}{dt} = x(1-x)(R-P) - \frac{P_{CD}}{2}\left((1-x)(R-S) + x(T-P)\right). \tag{7.5.8}$$

where P_{CD} is at the temporal equilibrium, i.e. $P_{CD} = -\frac{\rho}{1-\rho} + \sqrt{\left(\frac{\rho}{1-\rho}\right)^2 + \frac{4x(1-x)\rho}{1-\rho}}$ for all $0 < x < 1$ The boundary $x = 1$ is unstable but the boundary $x = 0$ is at least locally asymptotically stable since $d(dx/dt)/dt|_{x=1} = T - R > 0$ and $d(dx/dt)/dt|_{x=0} = S - P < 0$. The interior equilibrium of Eq. (7.5.8) is the solution of equation

$$\begin{aligned} & x(1-x)(R-P) - \frac{P_{CD}}{2}\left((1-x)(R-S) + x(T-P)\right) = 0. \\ \Rightarrow \quad & -\frac{\rho}{1-\rho} + \sqrt{\left(\frac{\rho}{1-\rho}\right)^2 + 4x(1-x)\frac{\rho}{1-\rho}} = \frac{2x(1-x)(R-P)}{(1-x)(R-S)+x(T-P)} \\ \Rightarrow \quad & x^2 - x\left[1 - \alpha A(T-S)\right] + \alpha(P-S)(R-S). \end{aligned} \tag{7.5.9}$$

where

$$A = R - S - T + P, \tag{7.5.10}$$

$$\alpha = \frac{\rho/(1-\rho)}{(R-P)^2 + (\rho/(1-\rho))A^2}. \tag{7.5.11}$$

So, two possible interior equilibria can be given by

$$x_{1,2}^* = \frac{1 + \alpha A(T-S) \pm \sqrt{(1 + \alpha A(T-S))^2 - 4\alpha(P-S)(R-S)}}{2}. \tag{7.5.12}$$

Here $x_1^* = x_2^* = (1 + \alpha A(T-S))/2$ if $(1 + \alpha A(T-S))^2 - 4\alpha(P-S)(R-S) = 0$.

For the local stability of interior equilibrium, a straight forward calculation shows that (i) if only one interior equilibrium exists, then it must be unstable; and (ii) for the situation with two interior equilibria x_1^* and x_2^* (with $x_1^* > x_2^*$), x_1^* is locally asymptotically stable but x_2^* is unstable.

Ethics

All participants provided written informed consent after the nature and possible consequences of the studies were explained. All experimental methods were carried out in accordance with the approved guidelines. All experimental protocols were approved by the Ethics Review Committee of Institute of Zoology.

Acknowledgements

Funding: We thank all participants of this study. This study was supported by the National Science Foundation of China (Grant No. 1131270439, 11471311 and 11301032) and the Natural Sciences and Engineering Research Council of Canada (Individual Discovery Grant).

Author Contributions: The experiment was designed by B.-Y.Z., S.-J.F. and C.L. with advice from Y.T. Data processing was carried out by B.-Y.Z. and S.-J.F., with critical input from R.C. and Y.T. Theoretical analysis was performed by B.-Y.Z., C.L., X.-D.Z., Y.T. and R.C. Experiment platform was coded by J.-Z.B. The manuscript was drafted by S.-J.F. and Y.T. and revised by R.C. All authors gave final approval for publication.

Bibliography

[1] W.D. Hamilton, The genetical evolution of social behaviour. *Journal of Theoretical Biology* **7**, 1-16 (1964).

[2] R. Trivers, The evolution of reciprocal altruism. *Quarterly Review of Biology* **46**, 35-57 (1971).

[3] R. Axelrod, W.D. Hamilton, The evolution of cooperation. *Science* **211**, 1390-1396 (1981).

[4] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[5] M.A. Nowak, K. Sigmund, Evolution of indirect reciprocity. *Nature* **437**, 1291-1298 (2005).

[6] M.A. Nowak, R.M. May, Evolutionary games and spatial chaos. *Nature* **359**, 826-829 (1992).

[7] H. Ohtsuki, C. Hauert, E. Lieberman, M.A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502-505 (2006).

[8] A. Traulsen, M.A. Nowak, Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 10952-10955 (2006).

[9] M.A. Nowak, Five rules for the evolution of cooperation. *Science* **314**, 1560-1563 (2006a).

[10] L.A. Dugatkin, *Cooperation Among Animals: An Evolutionary Perspective* (Oxford University Press, 1997).

[11] M.A. Nowak, *Evolutionary Dynamics* (Belknap Press, 2006b).

[12] K. Sigmund, *The Calculus of Selfishness* (Princeton University Press, 2010).

[13] R. Axelrod, D. Dion, The further evolution of cooperation. *Science* **242**, 1385-1390 (1988).

[14] I. Eshel, L.L. Cavalli-Sforza, Assortment of encounters and evolution of cooperativeness. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 1331-1335 (1982).

[15] L.L. Cavalli-Sforza, M.W. Feldman, Paradox of the evolution of communication and of social interactivity. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 2017-2021 (1983).

[16] J.A. Fletcher, M. Doebeli, How altruism evolves: assortment and synergy. *Journal of Evolutionary Biology* **19**, 1389-1393 (2006).

[17] C. Taylor, M.A. Nowak, Evolutionary game dynamics with non-uniform interaction rates. *Theoretical Population Biology* **69**, 243-252 (2006).

[18] J.M. Pacheco, A. Traulsen, H. Ohtsuki, M.A. Nowak, Repeated games and direct reciprocity under active linking. *Journal of Theoretical Biology* **250**, 723-731 (2008).

[19] N. Hayashi, From TIT-for-TAT to OUT-for-TAT. *Social Theory Methods* **8**, 19-32 (1993).

[20] R. Schuessler, Exit threats and cooperation under anonymity. *Journal of Conflict Resolution* **33**, 728-749 (1989).

[21] J.M. Orbell, R.M. Dawes, Social welfare, cooperators' advantage, and the option of not playing the game. *American sociological review* **58**(6), 787-800 (1993).

[22] E. Hauk, Multiple prisoner's dilemma games with (out) an outside option: an experimental study. *Theory and Decision* **54**(3), 207-229 (2003).

[23] C.A. Aktipis, Know when to walk away: contingent movement and the evolution of cooperation. *Journal of Theoretical Biology* **231**, 249-260 (2004).

[24] T. Fujiwara-Greve, M. Okuno-Fujiwara, Voluntarily separable repeated prisoner's dilemma. *The Review of Economic Studies* **76**, 993-1021 (2009).

[25] S.S. Izquierdo, L.R. Izquierdo, F. Vega-Redondo, The option to leave: Conditional dissociation in the evolution of cooperation. *Journal of Theoretical Biology* **267**, 76-84 (2010).

[26] S.S. Izquierdo, L.R. Izquierdo, F. Vega-Redondo, Leave and let leave: A sufficient condition to explain the evolutionary emergence of cooperation. *Journal of Economic Dynamics and Control* **46**, 91-113 (2014).

[27] J. Hofbauer, K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, 1998).

[28] J. Maynard-Smith, *Evolution and the Theory of Games* (Cambridge University Press, 1982).

[29] T. Clutton-Brock, Breeding together: kin selection and mutualism in cooperative vertebrates. *Science* **296**, 69-72 (2002).

[30] F. Vesely, C.L. Yang, Breakup, secret handshake and neutral stability in repeated Prisoner's Dilemma with option to leave: a note. Available at SSRN 2179126 (2012).

[31] C. Hauert, S. De Monte, J. Hofbauer, K. Sigmund, Volunteering as red queen mechanism for cooperation in public goods games. *Science* **296**, 1129-1132 (2002).

[32] D. Semmann, H.J. Krambeck, M. Milinski, Volunteering leads to rock-paper-scissors dynamics in a public goods game. *Nature* **425**, 390-393 (2003).

[33] F.C. Santos, J.M. Pacheco, T. Lenaerts, Cooperation prevails when individuals adjust their social ties. *PLoS Computational Biology* **2**(10), e140 (2006).

[34] B. Wu, D. Zhou, F. Fu, Q.-J. Luo, L. Wang, A. Traulsen, Evolution of cooperation on stochastic dynamical networks. *PLoS ONE* **5**(6), e11187 (2010).

[35] K. Fehl, D.J. van der Post, D. Semmann, Co-evolution of behaviour and social network structure promotes human cooperation. *Ecology letters* **14**(6), 546-551 (2011).

[36] D.G. Rand, S. Arbesman, N.A. Christakis, Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 19193-19198 (2011).

[37] J. Wang, S. Suri, D.J. Watts, Cooperation and assortativity with dynamic partner updating. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 14363-14368 (2012).

[38] P. Bednarik, K. Fehl, D. Semmann, Costs for switching partners reduce network dynamics but not cooperative behaviour. *Proceedings of the Royal Society B Biological Sciences* **281**, 20141661 (2014).

Chapter 8

Article 8

Effect of the opting-out strategy on conditions for selection to favor the evolution of cooperation in a finite population

Cong Li, Sabin Lessard

Département de mathématiques et de statistique
Université de Montréal, Montréal, Québec
H3C 3J7 CANADA

Information

This paper is currently under review. It was submitted to *Journal of Theoretical Biology* on 1 July 2020.

My contribution

Sabin Lessard initiated the project and provided some early calculations. I completed the theoretical framework and analysis and wrote the first draft of the paper.

Abstract

We consider a Prisoner's Dilemma (PD) that is repeated with positive probability $1-\rho$ only between cooperators as a result of an opting-out strategy practiced by all individuals. The population is made of N pairs of individuals and is updated at every time step by a birth-death event according to a Moran model. Assuming an intensity of selection of order $1/N$ and taking $2N^2$ birth-death events as unit of time, a diffusion approximation exhibiting two time scales, a fast one for pair frequencies and a slow one for cooperation (C) and defection (D) frequencies, is ascertained in the limit of a large population size. This diffusion approximation is applied to an an additive PD game, cooperation incurring a cost c and providing a benefit b to the opponent, and used to obtain the probability of ultimate fixation of a single C introduced as a single mutant in an all D population under selection and compare it to the probability under neutrality, which is $1/(2N)$, as well as the corresponding probability for a single D introduced as a single mutant in an all C population under selection. This gives conditions for cooperation to be favored by selection. We show that these conditions are satisfied when the benefit-to-cost ratio, b/c, exceeds some increasing function of ρ that is approximately given by $(1+\sqrt{\rho})/(1-\sqrt{\rho})$.

8.1 Introduction

In a two-player two-strategy game, known as the Prisoner's Dilemma (PD), in which cooperation and defection, denoted by C and D, respectively, are used by individuals in pairwise interactions, the payoffs are given by the entries of a 2×2 matrix

$$\begin{pmatrix} \pi_{CC} & \pi_{CD} \\ \pi_{DC} & \pi_{DD} \end{pmatrix} . \tag{8.1.1}$$

Here, π_{ij} represents the payoff to an individual using strategy i against an individual using strategy j where $i, j \in \{C, D\}$. In a PD game, cooperation against cooperation pays more than defection against defection, but less than defection against cooperation, while cooperation against defection pays the least. Thus, the entries of the payoff matrix satisfy the inequalities $\pi_{DC} > \pi_{CC} > \pi_{DD} > \pi_{CD}$ (see, e.g., Poundstone, 1992; Nowak and Highfield, 2011). In particular, if cooperation and defection have additive effects on the payoff with cooperation incurring a cost c but providing a benefit b to the opponent, then we have $\pi_{CC} = b - c$, $\pi_{CD} = -c$, $\pi_{DC} = b$ and $\pi_{DD} = 0$. We call this case the additive PD game.

In a one-round PD game with defection paying more than cooperation against both defection and cooperation, defection is the only rational choice and the only Nash equilibrium (NE) (see, e.g., Hofbauer and Sigmund, 1998; Nowak, 2006). In a repeated PD game, however, with pairwise interactions repeated between the same players a random number of times that does not depend on the strategies in use, the tit-for-tat (TFT) strategy starting with cooperation becomes a Nash equilibrium against the always-defect (AllD) strategy if the number of repetitions of the game is large enough (Axelrod and Hamilton, 1981; Axelrod, 1984). This can be seen as an effect of direct reciprocity (Trivers, 1971), since TFT against TFT leads to reciprocal cooperation and TFT or AllD against AllD to reciprocal defection at least after the first round.

On the other hand, if an individual can choose to repeat or not an interaction with an opponent in an iterated PD game, the rational choice is to repeat the interaction with a cooperator as long as possible, but end it with a defector as soon as possible. This is known as the opting-out or out-for-tat (OFT) strategy (Hayashi, 1993; Schuessler, 1989; Aktipis, 2004; Fujiwara-Greve and Okuno-Fujiwara, 2009; Izquierdo *et al.*, 2010, 2014). As a result, an interaction between two cooperators may be continued with some probability, while an interaction between two defectors or between one defector and one cooperator may never be repeated. This can be seen as a mechanism that creates direct reciprocity by which the evolution and maintenance of cooperation can be favored by selection (Zhang *et al.*, 2016; Zheng *et al.*, 2017; Kurokawa, 2019).

The opting-out strategy is akin to assortment of cooperative acts in social space, where cooperative behaviour is a repeatable trait of individuals, and cooperative individuals associate and interact with each other disproportionately more than with defectors (see, e.g., Eshel and Cavalli-Sforza, 1982). There is some evidence that individuals from a range of species show stability in their level of cooperativeness (Bergmüller *et al.*, 2010) and that animal social network structures may show significant within-population heterogeneity in social tie strengths (Krause *et al.* 2015). Recent empirical investigations in wild Trinidadian guppies (*Poecilia reticulata*) support the hypothesis that assortment by repeatable cooperativeness may be an important feature for the evolution and persistence of non-kin cooperation in real-world populations (Brask *et al.* 2019).

In this paper, we will consider a Prisoner's Dilemma (PD) that is repeated with positive probability $1-\rho$ only between cooperators as a result of an opting-out strategy practiced by all individuals. The population will be assumed to be made of N pairs of individuals and be updated at every time step by a birth-death event according to a Moran model. With appropriate scalings of the intensity of selection and time with respect to the population size, we will establish a diffusion approximation in the limit of a large population that shows that two time scales come into play, a fast one for pair frequencies and a slow one for C and D frequencies. This diffusion approximation will be applied to an additive PD game, cooperation incurring a cost c and providing a benefit b to the opponent, and used to obtain the probability of ultimate fixation of a single C introduced as a single mutant in an all D population under selection and compare it to the probability under neutrality, which is $1/(2N)$, as well as the corresponding probability for a single D introduced as a single mutant in an all C population under selection. This will provide conditions for cooperation to be favored by selection.

8.2 The Model

Consider a population of N pairs of interacting individuals in which each individual is either a cooperator, C, or a defector, D. The population state and its changes from time t to time $t + \Delta t$, a time interval of length $\Delta t = 1/(2N^2)$, are represented in Figure 8.1.

At time t, the number of CC pairs in the population is NP_{CC}, while the number of CD pairs is NP_{CD} and the number of DD pairs NP_{DD}. Then, $x = P_{CC} + P_{CD}/2$ is the frequency of C in the population, and $1 - x$ the frequency of D.

Suppose that all individuals in the population adopt the opting-out strategy so that only the individuals paired with a C partner are interested in continuing the interaction in the time interval $[t, t + \Delta t]$. As a result, all CD or DD pairs break apart, while each CC pair breaks apart with some probability ρ and, therefore, stays unbroken with probability $1 - \rho$. The parameter ρ is assumed to be a positive constant. Then, the number of free D individuals is

$$N_D = NP_{CD} + 2NP_{DD} = 2N(1 - x), \tag{8.2.1}$$

while the number of free C individuals is

$$N_C = NP_{CD} + 2R = 2R + 2N(x - P_{CC}), \tag{8.2.2}$$

where R stands for the number of broken CC pairs. This number is a random variable that follows a binomial distribution with parameters NP_{CC} and ρ.

Now assume that all free individuals form new pairs at random. The number of these is

$$\frac{N_C + N_D}{2} = NP_{CD} + NP_{DD} + R, \tag{8.2.3}$$

while the conditional expected frequencies of CC, CD and DD among these are

$$\frac{N_C(N_C-1)}{(N_C+N_D)(N_C+N_D-1)}, \quad \frac{2N_CN_D}{(N_C+N_D)(N_C+N_D-1)}, \quad \frac{N_D(N_D-1)}{(N_C+N_D)(N_C+N_D-1)}, \tag{8.2.4}$$

respectively. Besides, there are $N_{CC} = NP_{CC} - R$ unbroken CC pairs.

Let the random variables q_{CC}, q_{CD} and q_{DD} represent the frequencies of CC, CD and DD in the set made of all new pairs and all unbroken CC pairs. Note that $q_{CC}+q_{CD}/2 = x$ and $q_{DD}+q_{CD}/2 = 1-x$, which means that the frequencies of C and D in the population are unchanged. On the other hand, the expected values of q_{CC}, q_{CD} and q_{DD} are given by

$$\mathbf{E}(q_{CC}) = 2x - 1 + \frac{(1-x)^2}{1-(1-\rho)P_{CC}} + O(N^{-1/2}), \tag{8.2.5a}$$

$$\mathbf{E}(q_{CD}) = 2(1-x) - \frac{2(1-x)^2}{1-(1-\rho)P_{CC}} + O(N^{-1/2}), \tag{8.2.5b}$$

$$\mathbf{E}(q_{CC}) = \frac{(1-x)^2}{1-(1-\rho)P_{CC}} + O(N^{-1/2}), \tag{8.2.5c}$$

while their variances are all of order N^{-1}, that is,

$$\mathbf{Var}(q_{CC}) = O(N^{-1}), \tag{8.2.6a}$$

$$\mathbf{Var}(q_{CD}) = O(N^{-1}), \tag{8.2.6b}$$

$$\mathbf{Var}(q_{DD}) = O(N^{-1}). \tag{8.2.6c}$$

These results are shown in Appendix A.

The update of the population at the end of the time interval $[t, t+\Delta t]$ is obtained by a birth-death event according to a Moran model in a context of evolutionary game theory (see, e.g., Hofbauer and Sigmund, 1998; Ewens, 2004; Nowak *et al.* 2004; Ohtsuki *et al.* 2006). One individual is chosen with probability proportional to fitness to produce an

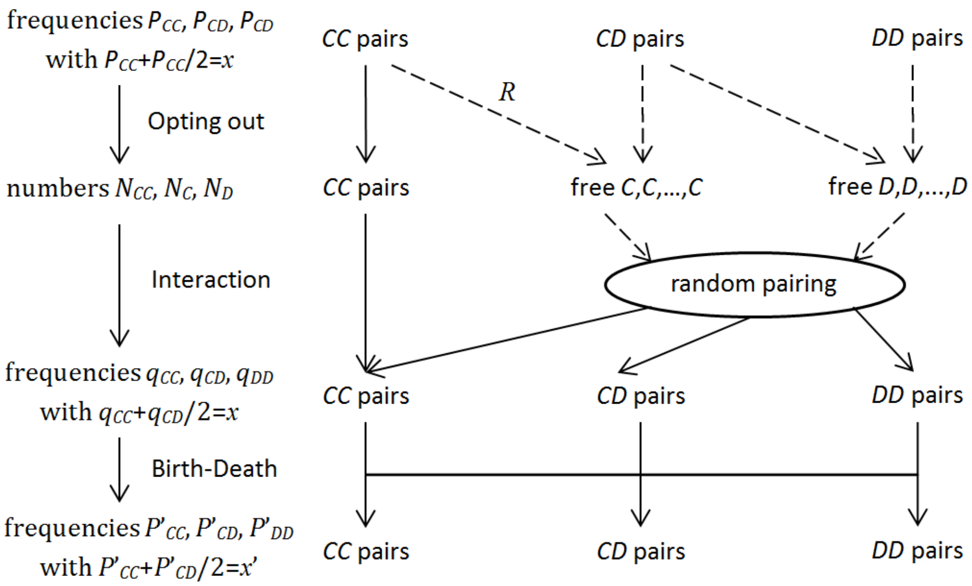


Figure 8.1: *Changes of the population state in a time interval* $[t, t+\Delta t]$.

offspring identical to itself and one individual is chosen at random to be replaced by the offspring.

Here, the fitness of an i-strategist in interaction with a j-strategist is given in the form

$$w_{ij} = 1 + s\pi_{ij}, \tag{8.2.7}$$

where 1 stands for a baseline fitness, $s = \sigma N^{-1}$ represents an intensity of selection, and π_{ij} denotes the payoff to i against j, for $i, j \in \{C, D\}$. Therefore, the offspring produced is a cooperator with conditional probability

$$\begin{aligned} Pr(C) &= \frac{2q_{CC}w_{CC} + q_{CD}w_{CD}}{2q_{CC}w_{CC} + q_{CD}w_{CD} + q_{CD}w_{DC} + 2q_{DD}w_{DD}} \\ &= x + \frac{\sigma}{N}\left((1-x)A - xB\right) + o(N^{-1}), \end{aligned} \tag{8.2.8}$$

and a defector with conditional probability

$$\begin{aligned} Pr(D) &= \frac{2q_{DD}w_{DD} + q_{CD}w_{DC}}{2q_{CC}w_{CC} + q_{CD}w_{CD} + q_{CD}w_{DC} + 2q_{DD}w_{DD}} \\ &= 1 - x - \frac{\sigma}{N}\left((1-x)A - xB\right) + o(N^{-1}), \end{aligned} \tag{8.2.9}$$

where

$$A = q_{CC}\pi_{CC} + \frac{q_{CD}}{2}\pi_{CD} \tag{8.2.10}$$

and

$$B = q_{DD}\pi_{DD} + \frac{q_{CD}}{2}\pi_{DC}. \tag{8.2.11}$$

Note that $Pr(C)$ and $Pr(D)$ are both random variables whose main terms are linear functions of q_{CC}, q_{CD} and q_{DD}.

On the other hand, the offspring produced replaces a cooperator with probability x, and a defector with probability $1-x$. Actually, it replaces a cooperator in a CC pair with probability q_{CC} or a CD pair with probability $q_{CD}/2$, while it replaces a defector in a DD pair with probability q_{DD} or a CD pair with probability $q_{CD}/2$.

Following the replacement of an individual by the offspring, the frequencies of CC, CD and DD among the N pairs are denoted by P'_{CC}, P'_{CD} and P'_{DD} respectively, and the frequency of C in these pairs is $P'_{CC} + P'_{CD}/2 = x'$. This gives the population state at the beginning of the next time interval which corresponds to time $t + \Delta t = t + 1/(2N^2)$ with $2N^2$ time intervals as unit of time. Note that P_{CD} and P_{DD} can be expressed in terms of x and P_{CC}, so that x and P_{CC} can be used to describe the population state.

8.3 Diffusion Approximation

Let $\Delta x = x' - x$ and $\Delta P_{CC} = P'_{CC} - P_{CC}$ be the changes in the frequencies of C and CC, respectively, from time t to time $t + \Delta t$ with $\Delta t = 1/(2N^2)$. Given these frequencies at time t, the conditional first, second and fourth moments of Δx are approximated as (see Appendix B for details)

$$\mathbf{E}(\Delta x) = \frac{1}{2N^2} m(x, P_{CC}) + o(N^{-2}) \ , \tag{8.3.1}$$

$$\mathbf{E}((\Delta x)^2) = \frac{1}{2N^2} v(x) + o(N^{-2}) \tag{8.3.2}$$

and

$$\mathbf{E}((\Delta x)^4) = o(N^{-3}) , \tag{8.3.3}$$

respectively, where

$$m(x, P_{CC}) = \sigma \mathbf{E}\left((1-x)A - xB\right) \tag{8.3.4}$$

and

$$v(x) = x(1-x). \tag{8.3.5}$$

Moreover, we have

$$\mathbf{E}(\Delta P_{CC}) = \frac{(x - P_{CC})^2 - \rho P_{CC}(1 - 2x + P_{CC})}{1 - (1-\rho)P_{CC}} + O(N^{-1/2}) \tag{8.3.6}$$

and

$$\mathbf{Var}(\Delta P_{CC}) = O(N^{-1}) \tag{8.3.7}$$

for the conditional mean and variance of ΔP_{CC}. On the other hand, in an infinite population in the absence of selection, the frequency of C remains constant while the frequency of CC converges uniformly to an equilibrium value P^*_{CC} in $[0,1]$. This equilibrium value is obtained by solving the equation $\mathbf{E}(\Delta P_{CC}) = 0$, which gives

$$P^*_{CC} = x + \frac{\rho}{2(1-\rho)} - \frac{\sqrt{\rho^2 + 4x(1-x)\rho(1-\rho)}}{2(1-\rho)}, \tag{8.3.8}$$

where x is the frequency of C (see Appendix C for details).

The conditions (8.3.1), (8.3.2), (8.3.3), (8.3.6), (8.3.7) and (8.3.8) show that there are two time scales at work in the discrete-time Markov chain for the population state, the variable P_{CC} changing more rapidly than the variable x. Moreover, as $N \to \infty$, these conditions ascertain that the Markov chain converges to a diffusion approximation with $m(x) = m(x, P^*_{CC})$ as drift function, and $v(x) = x(1-x)$ as diffusion function (Ethier and Nagylaki 1980).

Using (8.2.5) with $P_{CC} = P^*_{CC}$ and the equality (see (8.6.34) in Appendix C)

$$\frac{(1-x)^2}{1-(1-\rho)P^*_{CC}} = P^*_{CC} - 2x + 1 \tag{8.3.9}$$

leads to

$$\mathbf{E}(q_{CC}) = P^*_{CC} + O(N^{-1/2}), \tag{8.3.10a}$$

$$\mathbf{E}(q_{CD}) = 2x - 2P^*_{CC} + O(N^{-1/2}), \tag{8.3.10b}$$

$$\mathbf{E}(q_{DD}) = 1 - 2x + P^*_{CC} + O(N^{-1/2}). \tag{8.3.10c}$$

Let us summarize.

Result 8.1: *Consider a PD game with payoff matrix (8.1.1) for N pairs of individuals so that, as a result of opting-out from one round to the next, all pairs break apart to form new pairs at random but a random proportion of CC pairs whose mean is* $1-\rho < 1$. *Assume one birth-death event at the end of each round with the probability of giving birth proportional to an affine function of payoff with coefficient* σ/N *and the probability of dying given by* $1/(2N)$. *Taking* $2N^2$ *birth-death events as unit of time and letting* $N \to \infty$, *the Markov chain of the frequency of C converges to a diffusion with* $v(x) = x(1-x)$ *as diffusion function and*

$$m(x) = \sigma\left(x(1-x)(\pi_{CC} - \pi_{DD}) - (x - P^*_{CC})\left((1-x)(\pi_{CC} - \pi_{CD}) + x(\pi_{DC} - \pi_{DD})\right)\right) \tag{8.3.11}$$

as drift function, where P^*_{CC} *is given by (8.3.8).*

In the diffusion approximation, it is known (see, e.g., Kimura, 1964; Risken, 1992; Ewens, 2004) that the probability density function of C evaluated at x at time $t \geq 0$ given a value p at time 0, denoted by $f(x,p,t)$, satisfies the forward Kolmogorov (Fokker-Planck) equation

$$\frac{\partial f(x,p,t)}{\partial t} = -\frac{\partial}{\partial x}\left[m(x)f(x,p,t)\right] + \frac{\partial^2}{\partial x^2}\left[\frac{v(x)f(x,p,t)}{2}\right], \tag{8.3.12}$$

as well as the backward Kolmogorov equation

$$\frac{\partial f(x;p,t)}{\partial t} = m(p)\frac{\partial f(x;p,t)}{\partial p} + \frac{v(p)}{2}\frac{\partial^2 f(x;p,t)}{\partial p^2}. \tag{8.3.13}$$

In the case at hand with no mutation, the two boundaries $x = 0$ and $x = 1$ are absorbing states.

Moreover, if $u(p, t)$ denotes the probability that C is fixed by time $t \geq 0$, so that $x(t) = 1$ given an initial frequency $x(0) = p$, then it is known that this fixation probability satisfies the backward Kolmogorov equation, that is,

$$\frac{\partial u(p,t)}{\partial t} = m(p)\frac{\partial u(p,t)}{\partial p} + \frac{v(p)}{2}\frac{\partial^2 u(p,t)}{\partial p^2}, \tag{8.3.14}$$

with the boundary conditions $u(0, t) = 0$ and $u(1, t) = 1$. By letting $t \to \infty$, the limit

$$u(p) \equiv \lim_{t\to\infty} u(p,t) \tag{8.3.15}$$

represents the probability of ultimate fixation of C given an initial frequency $x(0) = p$. As $t \to \infty$, the left-hand side in (8.3.14) tends to 0 so that we have

$$0 = m(p)\frac{du(p)}{dp} + \frac{v(p)}{2}\frac{d^2u(p)}{dp^2} \tag{8.3.16}$$

with the boundary conditions $u(0) = 0$ and $u(1) = 1$. The solution of this ordinary differential equation is known to be (see, e.g., Ewens, 2004)

$$u(p) = \frac{\int_0^p \psi(y)dy}{\int_0^1 \psi(y)dy}\ , \tag{8.3.17}$$

where

$$\psi(y) = \exp\left(-2\int_0^y \frac{m(x)}{v(x)}dx\right)\ . \tag{8.3.18}$$

Note that the probability of ultimate fixation of D is given by $1 - u(p)$, since there is ultimate fixation of C or D with probability 1.

8.4 Additive PD game

Consider an additive Prisoner's Dilemma (PD) where a cooperator pays a fixed cost $c > 0$ while its partner receives a fixed benefit $b > c$. The payoff matrix (8.1.1) takes the form

$$\begin{pmatrix} \pi_{CC} & \pi_{CD} \\ \pi_{DC} & \pi_{DD} \end{pmatrix} = \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}. \tag{8.4.1}$$

Substituting the above payoffs into $m(x)$ in Result 8.1 yields

$$\begin{aligned} m(x) &= \sigma\left(x(1-x)(b-c) - b(x - P^*_{CC})\right) \\ &= \sigma\left[x(1-x)(b-c) - \frac{bf(x,\rho)}{2(1-\rho)}\right], \end{aligned} \tag{8.4.2}$$

where

$$f(x,\rho) = \sqrt{\rho^2 + 4x(1-x)\rho(1-\rho)} - \rho. \tag{8.4.3}$$

This function defines a concave surface on the domain $[0,1]\times[0,1]$ with $f(x,\rho) > 0$ inside this domain and $f(x,\rho) = 0$ on its boundary (see Figure 8.2a).

Now, let $F_C = u((2N)^{-1})$ be the probability of ultimate fixation of C introduced as a single mutant in an all D population of size $2N$. The corresponding fixation probability for a single D introduced in an all C population is $F_D = 1 - u(1-(2N)^{-1})$. The evolution of cooperation is said to be favored by selection if $F_C > (2N)^{-1}$, where $(2N)^{-1}$ is the fixation probability under neutrality. Similarly, the evolution of defection is said to be unfavored by selection if $F_D < (2N)^{-1}$. On the other hand, the evolution of cooperation is said to be more favored by selection than the evolution of defection if $F_C > F_D$. Finally, if the three conditions are simultaneously satisfied, that is, $F_C > (2N)^{-1} > F_D$, then the evolution of cooperation is said to be fully favored by selection (Nowak *et al.*, 2004; Li and Lessard, 2020).

When the population size $2N$ is big enough, the conditions to have $F_C > (2N)^{-1}$, $F_D < (2N)^{-1}$, $F_C > F_D$ take the form

$$\int_0^1 \psi(y)dy < 1, \tag{8.4.4a}$$

$$\int_0^1 \psi(y)dy > \psi(1), \tag{8.4.4b}$$

$$\psi(1) < 1, \tag{8.4.4c}$$

respectively (Li and Lessard, 2020). Here, we have

$$\psi(y) = \exp\left(-2\int_0^y g(x)dx\right), \tag{8.4.5}$$

where

$$\begin{aligned} g(x) = \frac{m(x)}{v(x)} &= \sigma c \left[(b/c - 1) - \frac{(b/c) f(x,\rho)}{2(1-\rho)x(1-x)} \right] \\ &= \sigma c \left[r \left(1 - \frac{2}{\sqrt{1+4ax(1-x)}+1} \right) - 1 \right], \end{aligned} \tag{8.4.6}$$

with the notation $r = b/c > 1$ for the benefit-to-cost ratio and $a = (1/\rho) - 1 \geq 0$ for the expected number of times that each CC pair continues to interact.

In the extreme case $a = 0$ (or $\rho = 1$ which means no repeated interactions between cooperators), we have $g(x) = -\sigma c < 0$ which implies that $\psi(y)$ is a strictly increasing function of $y \in [0,1]$ with $\psi(0) = 1$. All conditions in (8.4.4) are not satisfied. Cooperation is never favored by selection, while defection always is. This is exactly the case of the classic PD game. On the other hand, if $a \to \infty$ (or $\rho = 0$ which means permanent CC pairs), we have $g(x) = \sigma(b - c) > 0$ which implies that $\psi(y)$ is a strictly decreasing function of $y \in [0,1]$ with $\psi(0) = 1$. All conditions in (8.4.4) are satisfied. Cooperation is fully favored by selection. This is easy to understand since, in this case, CC pairs never break apart and their number can only increase.

Analogously, in the extreme case $r = 1$, we get $-\sigma c \leq g(x) < 0$ for $x \in [0,1]$ and, therefore, $\psi(y)$ is a strictly increasing function of $y \in [0,1]$ with $\psi(0) = 1$, which implies that all conditions in (8.4.4) are not satisfied and cooperation can never be favored by selection. In this case, the cooperators pay as much as they give and the game is actually no longer a PD game. On the other hand, if $r \to \infty$, for any given $a > 0$, there exists $r > 0$ such that $g(x) > 0$ for $x \in [0,1]$ which implies that $\psi(y)$ is a strictly decreasing function of $y \in [0,1]$ with $\psi(0) = 1$. In this case, conditions in (8.4.4) are all satisfied and cooperation is fully favored by selection.

In the general case $0 < \rho < 1$ and $r > 1$, the expression of $g(x)$ in (8.4.6) shows that $\partial g(x)/\partial a > 0$ and $\partial g(x)/\partial r > 0$. This gives that $\psi(y)$ is a strictly decreasing function of a and r for every $y \in (0,1]$. This leads to the following conclusion.

Result 8.2: *Consider an additive PD game with payoff matrix (8.4.1) in the frame-*

work of Result 8.1 with $0 < \rho < 1$. *In a large enough population, increasing the value of* $r = b/c > 1$ *or* $a = 1/\rho - 1 > 0$ *(or decreasing the value of* ρ*) increases (decreases, respectively) the probability of ultimate fixation of cooperation (defection, respectively) introduced as a single mutant in an all defecting (cooperating, respectively) population,* F_C *(*F_D*, respectively).*

The proof of this result is straightforward by using the approximations

$$F_C \approx \left(2N \int_0^1 \exp\left(-2\int_0^y g(x)dx\right) dy\right)^{-1}, \tag{8.4.7a}$$

$$F_D \approx \left(2N \int_0^1 \exp\left(2\int_y^1 g(x)dx\right) dy\right)^{-1}, \tag{8.4.7b}$$

for N large enough.

Moreover, since $g(x)$ is a symmetric function, that is, $g(x) = g(1-x)$ for $x \in [0,1]$. When $F_C = F_D$, that is, $\psi(1) = 1$, we have $\int_0^1 g(x)dx = 0$ from which

$$\begin{aligned}\int_0^{1-y} g(x)dx &= \int_0^1 g(x)dx - \int_{1-y}^1 g(x)dx \\ &= -\int_y^0 g(1-x)d(1-x) \\ &= -\int_0^y g(1-x)dx \\ &= -\int_0^y g(x)dx.\end{aligned} \tag{8.4.8}$$

In this case, we have

$$\begin{aligned}\int_0^1 \psi(y)dy = \int_0^{\frac{1}{2}} \psi(y)dy + \int_{\frac{1}{2}}^1 \psi(y)dy &= \int_0^{\frac{1}{2}} \big(\psi(y) + \psi(1-y)\big)dy \\ &= \int_0^{\frac{1}{2}} \left(\exp\left(-2\int_0^y g(x)dx\right) + \exp\left(-2\int_0^{1-y} g(x)dx\right)\right) dy \\ &= \int_0^{\frac{1}{2}} \left(\exp\left(-2\int_0^y g(x)dx\right) + \exp\left(2\int_0^y g(x)dx\right)\right) dy \\ &\geq 2\int_0^{\frac{1}{2}} dy = 1,\end{aligned} \tag{8.4.9}$$

with an equality if and only if $g(x) \equiv 0$ on $[0,1]$, which means $\sigma = 0$ (no selection). Otherwise, $\int_0^1 \psi(y)dy > 1 = \psi(1)$, which means that $F_C = F_D < (2N)^{-1}$ owing to

(8.4.4a,b,c). From the previous analysis in the extreme cases and Result 8.2, we know that increasing the value of a from 0 to ∞ or r from 1 to ∞, F_C will increase F_C from a value smaller than $(2N)^{-1}$ to a value larger than $(2N)^{-1}$, and will decrease F_D in the opposite direction. But when F_C and F_D are equal, their values are less than $(2N)^{-1}$. This implies that F_D crosses the value $(2N)^{-1}$ first, then equals F_C and finally F_C crosses the value $(2N)^{-1}$. Thus we get the following corollary of Result 8.2.

Result 8.3: *In the setting of Result 8.2, as the value of r or a increases, the conditions (8.4.4b), (8.4.4c) and (8.4.4a) for $F_D < (2N)^{-1}$, $F_C > F_D$ and $F_C > (2N)^{-1}$, respectively, are satisfied in this order. In particular, when cooperation is favored by selection, it is necessarily fully favored by selection.*

In order to get explicit conditions on the parameters of the model for cooperation to be favored by selection, we use the inequalities (see Appendix D for details)

$$4x(1-x)(\sqrt{\rho}-\rho) \le f(x,\rho) \le \sqrt{4x(1-x)}(\sqrt{\rho}-\rho), \tag{8.4.10}$$

where the lower bound is the limit of $f(x,\rho)$ as $\rho \to 0$ and the upper bound the limit of $f(x,\rho)$ as $\rho \to 1$. Panels (a), (b) and (c) in Figure 8.2 illustrate the surfaces determined by $f(x,\rho)$, $4x(1-x)(\sqrt{\rho}-\rho)$ and $\sqrt{4x(1-x)}(\sqrt{\rho}-\rho)$, respectively. Panels (d) and (e) show the transverse sections where $f(x,\rho)$ approaches the upper bound when ρ is close to 0 and the lower bound when ρ is close to 1, respectively.

Now, substituting $g(x)$ given in (8.4.6) in the expression of $\psi(y)$ given in (8.4.5) yields

$$\psi(y) = \exp\left(\sigma c\left(-2(r-1)y + \frac{r}{1-\rho}\int_0^y \frac{f(x,\rho)}{x(1-x)}dx\right)\right). \tag{8.4.11}$$

Panel (f) in Figure 8.2 shows that the lower bound of $f(x,\rho)$ gives a good approximation of the integral in (8.4.11) for $y=1$. Using this lower bound in this equation yields the approximation

$$\psi(y) \approx \exp\left(-2\sigma c\left(r-1-r\frac{2\sqrt{\rho}}{1+\sqrt{\rho}}\right)y\right), \tag{8.4.12}$$

which is a monotonic function of y starting with the value 1 at $y=0$. Using this approximation for $\psi(y)$ in (8.4.4), we see that this approximation would have to be a strictly

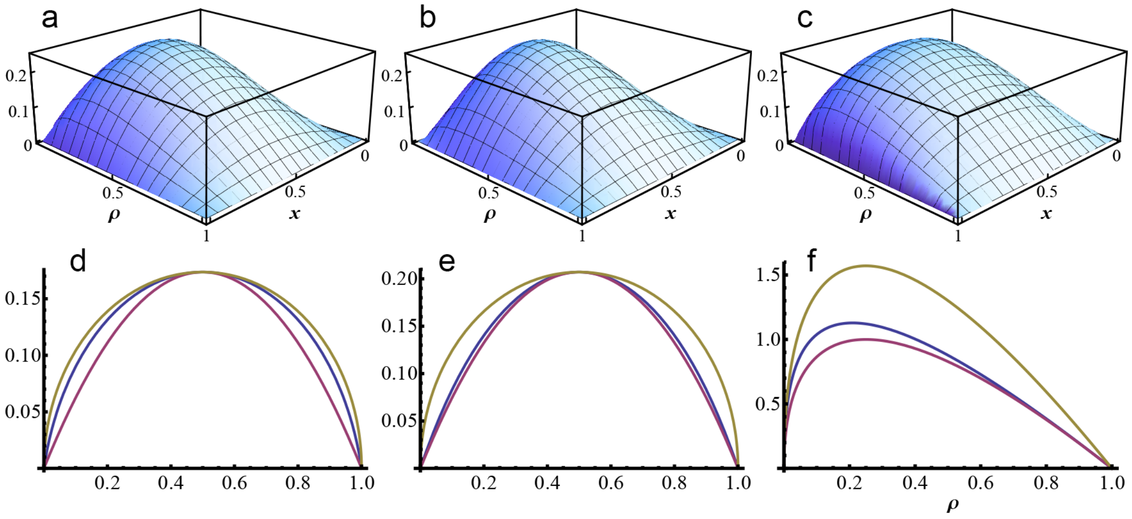


Figure 8.2: *Function* $f(x,\rho)$*.* Panels (a), (b) and (c) show the surfaces determined by the function $f(x,\rho)$, the lower bound $4x(1-x)(\sqrt{\rho}-\rho)$ and the upper bound $\sqrt{4x(1-x)}(\sqrt{\rho}-\rho)$, respectively. Panels (d) and (e) show the transverse sections for $\rho = 0.05$ and $\rho = 0.5$. Panel (f) shows $\int_0^1 \frac{f(x,\rho)}{x(1-x)}dx$ which comes into play in $\psi(1)$. The blue, red and yellow curves in panels (d), (e) and (f) stand for $f(x,\rho)$, $4x(1-x)(\sqrt{\rho}-\rho)$ and $\sqrt{4x(1-x)}(\sqrt{\rho}-\rho)$, respectively.

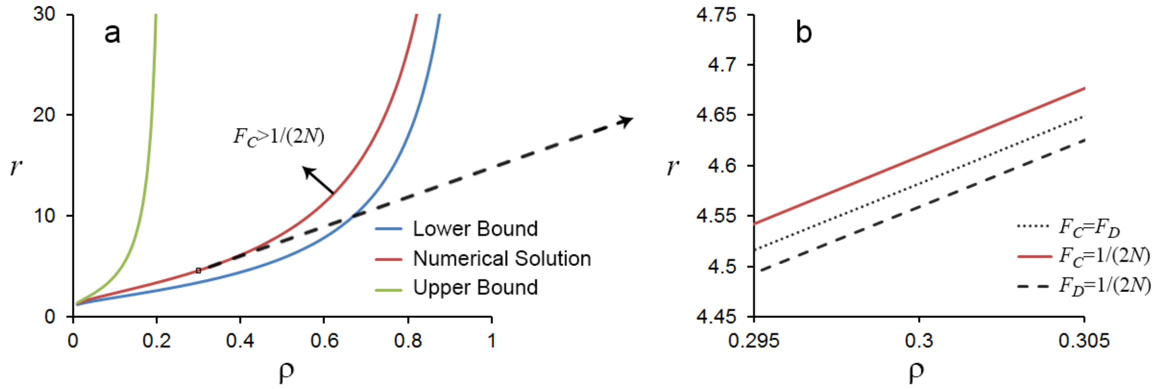


Figure 8.3: *Conditions for cooperation to be favored by selection in the additive PD game in the case* $c = 1$ *and* $\sigma = 1$*.* In panel (a), the red curve stands for the exact numerical solutions of $F_D = (2N)^{-1}$, $F_C = F_D$ and $F_C = (2N)^{-1}$, which are almost identical, while the blue and green curves stand for approximations obtained by using the lower and upper bounds of $f(x, \rho)$ plotted in Figure 8.2. Panel (b) is a magnification of the three curves for a small region of the domain in Panel (a).

decreasing function for cooperation to be favored by selection in any sense, that is,

$$r > \frac{1+\sqrt{\rho}}{1-\sqrt{\rho}}, \tag{8.4.13}$$

which is equivalent to

$$\rho < \left(\frac{r-1}{r+1}\right)^2. \tag{8.4.14}$$

The right-hand side in (8.4.13) is a lower threshold value for r, or equivalently, the right-hand side in (8.4.14) an upper threshold value for ρ. Analogously, using the upper bound of $f(x,\rho)$ in (8.4.11) yields the approximation

$$\psi(y) \approx \exp\left(-2\sigma c\left((r-1)y - r\frac{2\sqrt{\rho}}{(1+\sqrt{\rho})}\arcsin(y)\right)\right). \tag{8.4.15}$$

Using this approximation in (8.4.4c), we conclude that selection would favor more cooperation than defection if

$$r > \frac{1+\sqrt{\rho}}{1+\sqrt{\rho}-\pi\sqrt{\rho}}, \tag{8.4.16}$$

or equivalently,

$$\rho < \left(\frac{r-1}{(\pi-1)r+1}\right)^2. \tag{8.4.17}$$

The right-hand side in (8.4.16) is an upper threshold value for r, while the right-hand side in (8.4.17) is a lower threshold value for ρ. Note that the right-hand side in (8.4.16) goes to $+\infty$ as $\rho \to (\pi-1)^{-2} \approx 0.218$. As shown in Figure 8.2f, the approximation of $f(x,\rho)$ by the upper bound is not that good unless ρ is small enough.

The values of r or ρ such that $F_D = (2N)^{-1}$, $F_C = F_D$ and $F_C = (2N)^{-1}$, respectively, are illustrated in Figure 8.3 in the case $c = 1$ and $\sigma = 1$. The relative positions of the three curves using the exact expression of $f(x,\rho)$ given in (8.4.3) are in agreement with Result 8.3 but very close to each other. These curves are compared to the two curves obtained by using the lower and upper bounds of $f(x,\rho)$ given in (8.4.10). These correspond to the boundaries of the regions defined by (8.4.13) and (8.4.16), respectively. It can be seen that the first one based on the lower bound gives a good approximation. The form of the curves suggests that the conditions $F_D < (2N)^{-1}$, $F_C > F_D$ and $F_C > (2N)^{-1}$, which are all in favor of the evolution of cooperation, are in the form r greater than some increasing function of ρ, and the approximation obtained from the lower bound of $f(x, \rho)$ that this function is close to the one given in (8.4.13).

8.5 Discussion

Direct reciprocity is one of the most important mechanisms that can promote the evolution of cooperation (Trivers, 1971; Axelrod and Hamilton, 1981; Axelrod, 1984). In a repeated Prisoner's Dilemma (PD) game, for instance, tit-for-tat (TFT) starting with cooperation is a form of reciprocity and known to be a Nash equilibrium against always-defect (AllD) if the number of repetitions of the game is large enough (Nowak *et al.*, 2004; Nowak and Sigmund 2007). In this paper, we have considered the case where the PD game can be repeated only if both players are willing to continue their interaction, which occurs only when both cooperate. This is a rational choice when all players practice what is known as the opting-out strategy. This creates a kind of assortment that benefits cooperation (C) over defection (D) and should promote its evolution (Hayashi, 1993; Schuessler, 1989; Aktipis, 2004; Fujiwara-Greve and Okuno-Fujiwara, 2009; Izquierdo *et al.*, 2010, 2014).

A theoretical analysis of the effects of the opting-out strategy on the dynamics of a PD game in an infinite population can be found in Zheng *et al.* (2017). This uses the replicator equation for the C and D frequencies under the assumption of instantaneous equilibrium pair frequencies. In this paper, we have considered the same question but in a finite population in discrete time and in the limit of a large population size. We have shown that the dynamics of the population state over successive birth-death events according to a Moran model is approximated by a continuous-time diffusion if the intensity of selection and time are appropriately scaled with respect to the population size (Result 8.1). This has been ascertained by verifying conditions given in Ethier and Nagylaki (1980) for Markov chains with two timescales, here a fast one for pair frequencies and a slow one for C and D frequencies. Note that the drift function in this diffusion approximation given by $m(x)$ in (8.4.2), where x is the frequency of C, corresponds to the growth rate of x given by the replicator equation in Zheng *et al.* (2017). In the diffusion approximation, however, the boundaries $x = 0$ and $x = 1$ are absorbing states and a stationary distribution with coexistence of C and D is precluded unless a certain level of mutation is introduced. This is not the case with the replicator equation for an infinite population.

Assuming an additive PD game with cooperation incurring a cost c and providing a benefit b to the opponent, we have shown that increasing the benefit-to-cost ratio, $r = b/c$, or the expected number of repetitions of the PD game for a CC pair, $a = 1/\rho - 1$, makes it easier for the evolution of cooperation to be favored by selection, or for the evolution of defection to be unfavored by selection, or for the evolution of cooperation to be more favored by selection than the evolution of defection (Result 8.2). Here, this is understood in the sense that the probability of ultimate fixation of C introduced as a single mutant in an all D population under selection exceeds what it would be under neutrality, which is given by its initial frequency, or that the probability of ultimate fixation of D introduced as a single mutant in an all C population under selection is less than its initial frequency, or that the former probability exceeds the latter. Note that the first condition is the most stringent one and the second condition the least stringent one (Result 8.3). Moreover, the three conditions take the form r greater than some increasing function of ρ that has been shown to be approximated by $(1 + \sqrt{\rho})/(1 - \sqrt{\rho})$ (see Figure 8.3).

The condition $r > (1+\sqrt{\rho})/(1-\sqrt{\rho})$ for selection to favor the evolution of cooperation in the case of an additive PD game with opting-out in a large finite population is equivalent to $\rho < (b-c)^2/(b+c)^2$. This happens to be the condition for the existence of an interior equilibrium in an infinite population (Zheng *et al.*, 2017). On the other hand, comparing to the situation with TFT against AllD in a repeated PD game, we note that the condition for selection to favor the evolution of TFT in a large finite population is $r > (1+2\rho)/(1-\rho)$

(Nowak *et al.*, 2004; Nowak and Sigmund, 2007). Since $(1+\sqrt{\rho})/(1-\sqrt{\rho}) > (1+2\rho)/(1-\rho)$ for $\rho \in (0,1)$, it happens that the condition for selection to favor the evolution of cooperation is more stringent with opting-out than without opting-out. This is somehow in agreement with experimental results (Zhang *et al.*, 2016) showing a higher level of cooperation in groups without the possibility of opting-out than in groups using opting-out. However, this does not necessarily mean that TFT is better than the opting-out strategy in promoting the evolution of cooperation. Monte Carlo simulations (Izquierdo *et al.*, 2010; Zheng *et al.*, 2017) have shown that opting-out can prevail more than TFT in populations with these two strategies in use. Thus, a theoretical analysis of a model involving both TFT and opting-out would be of interest for future works.

The opting-out strategy provides an opportunity not only for cooperators to find cooperative partners but also for defectors who have an even greater advantage to do so. Moreover, ending an interaction with someone might incur a cost since there is a risk of not finding a new partner in time. In our model, there is no cost for opting-out. A cost could affect cooperators and defectors to different degrees and, therefore, the level of cooperation reached in the population.

Finally, the work in this paper has focussed on a two-player game. Kurokawa (2019) has studied the effect of opting-out on a three-player game in an infinite population. Extensions to n-player public goods game would be of interest.

8.6 Appendix

8.6.1 Appendix A: Mean and variance of pair frequencies after re-pairing of free individuals

The equations (8.2.1) and (8.2.2) for the numbers of free C and D individuals can be written into the form

$$N_C = N\gamma, \tag{8.6.1a}$$

$$N_D = N\beta, \tag{8.6.1b}$$

$$N_C + N_D = 2N\alpha. \tag{8.6.1c}$$

where

$$\gamma = \frac{2R}{N} + P_{CD} \leq 2, \tag{8.6.2a}$$

$$\beta = P_{CD} + 2P_{DD} = 2(1-x) \leq 2, \tag{8.6.2b}$$

$$\alpha = \frac{\beta+\gamma}{2} = \frac{R}{N} + P_{CD} + P_{DD} \leq 1, \tag{8.6.2c}$$

with R being a binomial random variable with parameters NP_{CC} and ρ.

The number of new pairs formed at random by all free individual is $(N_C+N_D)/2 = \alpha N$. Besides, there are $N_{CC} = NP_{CC} - R = N - \alpha N$ unbroken CC pairs. The new pairs are obtained by pairing the $(2k-1)$-th and $2k$-th free individuals chosen at random without replacement for $k = 1, \ldots, \alpha N$. Let $X_{CC,k}$, $X_{CD,k}$ and $X_{DD,k}$ be the random variables that take the value 1 if the k-th new pair is of types CC, CD and DD, respectively, and 0 otherwise, for $k = 1, ..., \alpha N$. Then, the numbers of new CC, CD and DD pairs can be expressed as

$$Y_{CC} = \sum_{k=1}^{\alpha N} X_{CC,k}, \tag{8.6.3a}$$

$$Y_{CD} = \sum_{k=1}^{\alpha N} X_{CD,k}, \tag{8.6.3b}$$

$$Y_{DD} = \sum_{k=1}^{\alpha N} X_{DD,k}, \tag{8.6.3c}$$

respectively. By symmetry, the first and second conditional moments of $X_{CC,k}$, $X_{CD,k}$ and $X_{DD,k}$ are given by

$$\mathbb{E}(X_{CC,k}) = \mathbb{E}(X_{CC,k}^2) = \mathbb{E}(X_{CC,1}) = \frac{N_C}{2\alpha N}\frac{N_C - 1}{2\alpha N - 1}, \tag{8.6.4a}$$

$$\mathbb{E}(X_{CD,k}) = \mathbb{E}(X_{CD,k}^2) = \mathbb{E}(X_{CD,1}) = 2\,\frac{N_C}{2\alpha N}\frac{N_D}{2\alpha N - 1}, \tag{8.6.4b}$$

$$\mathbb{E}(X_{DD,k}) = \mathbb{E}(X_{DD,k}^2) = \mathbb{E}(X_{DD,1}) = \frac{N_D}{2\alpha N}\frac{N_D - 1}{2\alpha N - 1}, \tag{8.6.4c}$$

$$\mathbb{E}(X_{CC,k}X_{CC,l}) = \mathbb{E}(X_{CC,1}X_{CC,2}) = \frac{N_C}{2\alpha N}\frac{N_C - 1}{2\alpha N - 1}\frac{N_C - 2}{2\alpha N - 2}\frac{N_C - 3}{2\alpha N - 3}, \tag{8.6.4d}$$

$$\mathbb{E}(X_{CD,k}X_{CD,l}) = \mathbb{E}(X_{CD,1}X_{CD,2}) = 4\,\frac{N_C}{2\alpha N}\frac{N_D}{2\alpha N - 1}\frac{N_C - 1}{2\alpha N - 2}\frac{N_D - 1}{2\alpha N - 3}, \tag{8.6.4e}$$

$$\mathbb{E}(X_{DD,k}X_{DD,l}) = \mathbb{E}(X_{DD,1}X_{DD,2}) = \frac{N_D}{2\alpha N}\frac{N_D - 1}{2\alpha N - 1}\frac{N_D - 2}{2\alpha N - 2}\frac{N_D - 3}{2\alpha N - 3}. \tag{8.6.4f}$$

where $l \neq k$. In particular, this yields

$$\mathbb{E}(Y_{CC}) = \frac{N_C(N_C - 1)}{2(2\alpha N - 1)} = \frac{\gamma N(\gamma N - 1)}{2(2\alpha N - 1)}, \tag{8.6.5a}$$

$$\mathbb{E}(Y_{CD}) = \frac{N_C N_D}{2\alpha N - 1} = \frac{\gamma N \beta N}{2\alpha N - 1} = \frac{(2\alpha N - \beta N)\beta N}{2\alpha N - 1}, \tag{8.6.5b}$$

$$\mathbb{E}(Y_{DD}) = \frac{N_D(N_D - 1)}{2(2\alpha N - 1)} = \frac{\beta N(\beta N - 1)}{2(2\alpha N - 1)}. \tag{8.6.5c}$$

Now, let q_{CC}, q_{CD} and q_{DD} represent the random frequencies of CC, CD and DD among the αN new pairs and $N - \alpha N$ unbroken CC pairs. Their expected values are given by

$$\mathbf{E}(q_{DD}) = \mathbf{E}\left(\frac{Y_{DD}}{N}\right) = \frac{1}{2}\mathbf{E}\left(\frac{\beta^2 N - \beta}{2\alpha N - 1}\right), \tag{8.6.6a}$$

$$\mathbf{E}(q_{CD}) = \mathbf{E}\left(\frac{Y_{CD}}{N}\right) = \mathbf{E}\left(\frac{\beta(2\alpha N - \beta N)}{2\alpha N - 1}\right) = \beta - \mathbf{E}\left(\frac{\beta^2 N - \beta}{2\alpha N - 1}\right), \tag{8.6.6b}$$

$$\mathbf{E}(q_{CC}) = 1 - \mathbf{E}(q_{CD}) - \mathbf{E}(q_{DD}) = 1 - \beta + \frac{1}{2}\mathbf{E}\left(\frac{\beta^2 N - \beta}{2\alpha N - 1}\right). \tag{8.6.6c}$$

Note that

$$\frac{\beta^2 N - \beta}{2\alpha N - 1} = \beta \frac{\beta - 1/N}{\beta + \gamma - 1/N} \leq \beta \leq 2. \tag{8.6.7}$$

Moreover, the expected value of this random variable is given by

$$\mathbf{E}\left(\frac{\beta^2 N - \beta}{2\alpha N - 1}\right) = \frac{\beta^2}{2(\rho P_{CC} + P_{CD} + P_{DD})} + O(N^{-1/2}). \tag{8.6.8}$$

This is obviously true when $\beta = 0$. On the other hand, when $\beta = P_{CD} + 2P_{DD} > 0$, we have $P_{CD} \geq 1/N$ or $P_{DD} \geq 1/N$, from which $\alpha = R/N + P_{CD} + P_{DD} \geq P_{CD} + P_{DD} \geq 1/N$. In this case, we have

$$0 \leq \frac{\beta}{\alpha - 1/(2N)} \leq \frac{P_{CD} + 2P_{DD}}{P_{CD} + P_{DD} - 1/(2N)} \leq 2\frac{P_{CD} + P_{DD}}{P_{CD} + P_{DD} - 1/(2N)} \leq 4. \tag{8.6.9}$$

Moreover, $\beta = P_{CD} + 2P_{DD} \leq 2(\rho P_{CC} + P_{CD} + P_{DD})$. Using these inequalities and the Cauchy-Schwarz inequality yields

$$\begin{aligned}
&\left|\mathbf{E}\left(\frac{\beta^2 N - \beta}{2\alpha N - 1}\right) - \frac{\beta^2}{2(\rho P_{CC} + P_{CD} + P_{DD})}\right| \\
&\leq \frac{1}{2}\mathbf{E}\left(\left|\frac{\beta(\beta - 1/N)}{\alpha - 1/(2N)} - \frac{\beta^2}{\rho P_{CC} + P_{CD} + P_{DD}}\right|\right) \\
&= \frac{1}{2}\mathbf{E}\left(\frac{\beta}{\alpha - 1/(2N)}\left|\beta - \frac{1}{N} - \frac{\beta(\alpha - 1/(2N))}{\rho P_{CC} + P_{CD} + P_{DD}}\right|\right) \\
&\leq 2\mathbf{E}\left(\left|\frac{\beta(\rho P_{CC} - R/N + 1/(2N))}{\rho P_{CC} + P_{CD} + P_{DD}} - \frac{1}{N}\right|\right) \\
&\leq 2\mathbf{E}\left(\left|\frac{\beta(\rho P_{CC} - R/N + 1/(2N))}{\rho P_{CC} + P_{CD} + P_{DD}}\right| + \frac{1}{N}\right) \\
&\leq 4\mathbf{E}\left(\left|\rho P_{CC} - \frac{R}{N} + \frac{1}{2N}\right|\right) + \frac{2}{N} \\
&\leq \frac{4}{N}\mathbf{E}\left(|\rho N P_{CC} - R|\right) + \frac{4}{N} \\
&\leq \frac{4}{N}\left(\mathbf{E}\left((\rho N P_{CC} - R)^2\right)\right)^{1/2} + \frac{4}{N} \\
&= 4\left(\frac{P_{CC}\rho(1-\rho)}{N}\right)^{1/2} + \frac{4}{N} \\
&= O(N^{-1/2})
\end{aligned} \tag{8.6.10}$$

This proves Eq. (8.6.8). Using this equation and the equality $\beta^2 = 4(1-x)^2$ in (8.6.6) yield

$$\mathbf{E}(q_{DD}) = \frac{(1-x)^2}{1-(1-\rho)P_{CC}} + O(N^{-1/2}) \tag{8.6.11a}$$

$$\mathbf{E}(q_{CD}) = 2(1-x) - \frac{2(1-x)^2}{1-(1-\rho)P_{CC}} + O(N^{-1/2}), \tag{8.6.11b}$$

$$\mathbf{E}(q_{CC}) = 2x - 1 + \frac{(1-x)^2}{1-(1-\rho)P_{CC}} + O(N^{-1/2}). \tag{8.6.11c}$$

Moreover, using the inequality $2\alpha \geq \beta$, we have

$$\begin{aligned}
\mathbf{Var}(q_{DD}) &= \mathbf{Var}\left(\frac{Y_{DD}}{N}\right) \\
&= \frac{1}{N^2}\left(\mathbf{E}(Y_{DD}^2) - (\mathbf{E}(Y_{DD}))^2\right) \\
&= \frac{1}{N^2}\left(\mathbf{E}\left(\left(\sum_{k=1}^{\alpha N} X_{DD,k}\right)^2\right) - \left(\mathbf{E}\left(\sum_{k=1}^{\alpha N} X_{DD,k}\right)\right)^2\right) \\
&= \frac{1}{N^2}\mathbf{E}\left(\alpha N(\alpha N-1)\mathbb{E}(X_{DD,1}X_{DD,2}) + \alpha N\mathbb{E}(X_{DD,1}^2) - (\alpha N\mathbb{E}(X_{DD,1}))^2\right) \\
&= \frac{1}{N^2}\mathbf{E}\Big(\alpha N(\alpha N-1)\frac{\beta N}{2\alpha N}\frac{\beta N-1}{2\alpha N-1}\frac{\beta N-2}{2\alpha N-2}\frac{\beta N-3}{2\alpha N-3} \\
&\quad + \alpha N\frac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)} - \alpha^2 N^2\left(\frac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\right)^2\Big) \\
&= \mathbf{E}\left(\alpha^2\frac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\left(\frac{(\beta N-2)(\beta N-3)}{(2\alpha N-2)(2\alpha N-3)} - \frac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\right)\right) \\
&\quad + \mathbf{E}\left(\frac{\alpha}{N}\frac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\left(1 - \frac{(\beta N-2)(\beta N-3)}{(2\alpha N-2)(2\alpha N-3)}\right)\right) \\
&\leq \frac{1}{N}\mathbf{E}\left(\alpha\frac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\left(1 - \frac{(\beta N-2)(\beta N-3)}{(2\alpha N-2)(2\alpha N-3)}\right)\right) \\
&\leq \frac{1}{N},
\end{aligned} \tag{8.6.12}$$

which implies that $\mathbf{Var}(q_{DD}) = O(N^{-1})$. Analogously, using the inequality $2\alpha \geq \gamma$, we can ascertain that

$$\mathbf{Var}(q_{CD}) = \mathbf{Var}\left(\frac{Y_{CD}}{N}\right) = O(N^{-1}). \tag{8.6.13}$$

Since $q_{CC} + q_{CD} + q_{DD} = 1$, we have also

$$\mathbf{Var}(q_{CC}) = \mathbf{Var}(1 - q_{CD} - q_{DD}) \leq 2(\mathbf{Var}(q_{CD}) + \mathbf{Var}(q_{DD})) = O(N^{-1}). \tag{8.6.14}$$

8.6.2 Appendix B: Moments of changes in C and CC frequencies

From time t to time $t + \Delta t$, the frequencies of CC, CD and DD pairs go from P_{CC}, P_{CD} and P_{DD}, with $x = P_{CC} + P_{CD}/2$ as frequency of C, to P'_{CC}, P'_{CD} and P'_{DD}, with $x' = P'_{CC} + P'_{CD}/2$ as frequency of C, after random re-pairing of free individuals, weighted random sampling of an individual to produce an offspring, with weights given by fitness, and random sampling of an individual to be replaced by the offspring. The random frequencies of CC, CD and DD pairs after the first step are q_{CC}, q_{CD} and q_{DD} with $x = q_{CC} + q_{CD}/2$, while according to (8.2.8) and (8.2.9), the conditional probabilities to sample C and D at the second step are

$$Pr(C) = x + \frac{\sigma}{N}\left((1-x)A - xB\right) + o(N^{-1}) \tag{8.6.15}$$

and

$$Pr(D) = 1 - x - \frac{\sigma}{N}\left((1-x)A - xB\right) + o(N^{-1}), \tag{8.6.16}$$

respectively, where

$$A = q_{CC}\pi_{CC} + \frac{q_{CD}}{2}\pi_{CD} \tag{8.6.17}$$

and

$$B = q_{DD}\pi_{DD} + \frac{q_{CD}}{2}\pi_{DC}. \tag{8.6.18}$$

Obviously, the probabilities to sample C and D at the third step are x and $1-x$, respectively.

Change in C frequency

The change in the frequency of C from time t to time $t+\Delta t$, represented by $\Delta x = x' - x$, takes the values $1/(2N)$ and $-1/(2N)$ with conditional probabilities $(1-x)Pr(C)$ and

$xPr(D)$, respectively, and 0 otherwise. The expected value of this change is

$$\begin{aligned}\mathbf{E}(\Delta x) &= \mathbf{E}\left(\frac{1}{2N}(1-x)Pr(C) - \frac{1}{2N}xPr(D)\right)\\ &= \frac{\sigma}{2N^2}\mathbf{E}\left((1-x)A - xB\right) + o(N^{-2}).\end{aligned} \tag{8.6.19}$$

The second moment is given by

$$\begin{aligned}\mathbf{E}((\Delta x)^2) &= \mathbf{E}\left(\frac{1}{4N^2}(1-x)Pr(C) + \frac{1}{4N^2}xPr(D)\right)\\ &= \frac{1}{2N^2}x(1-x) + o(N^{-2}).\end{aligned} \tag{8.6.20}$$

As for the fourth conditional moment, we have

$$\mathbf{E}((\Delta x)^4) = \mathbf{E}\left(\frac{1}{16N^4}(1-x)Pr(C) + \frac{1}{16N^4}xPr(D)\right) = o(N^{-3}). \tag{8.6.21}$$

Change in CC frequency

The change in the frequency of CC pairs from time t to time $t + \Delta t$ is given by

$$\Delta P_{CC} = P'_{CC} - P_{CC} = (P'_{CC} - q_{CC}) + q_{CC} - P_{CC}. \tag{8.6.22}$$

Given q_{CC}, q_{CD} and q_{DD}, the difference $P'_{CC} - q_{CC}$ takes the values $1/N$ and $-1/N$ with conditional probabilities $(q_{CD}/2)Pr(C)$ and $q_{CC}Pr(D)$, respectively, and 0 otherwise. Therefore, The difference has an expected value $O(N^{-1})$ and a variance $O(N^{-2})$. Using this and the expression of the expected values and variances of q_{CC}, q_{CD} and q_{DD} given in Appendix A, we get

$$\begin{aligned}\mathbf{E}(\Delta P_{CC}) &= \mathbf{E}(P'_{CC} - q_{CC}) + \mathbf{E}(q_{CC}) - P_{CC}\\ &= 2x - 1 + \frac{(1-x)^2}{1-(1-\rho)P_{CC}} - P_{CC} + O(N^{-1/2})\\ &= \frac{(x-P_{CC})^2 - \rho P_{CC}(1-2x+P_{CC})}{1-(1-\rho)P_{CC}} + O(N^{-1/2})\end{aligned} \tag{8.6.23}$$

and

$$\mathbf{Var}(\Delta P_{CC}) = \mathbf{Var}((P'_{CC} - q_{CC}) + q_{CC}) \leq 2(\mathbf{Var}(P'_{CC} - q_{CC}) + \mathbf{Var}(q_{CC})) = O(N^{-1}). \tag{8.6.24}$$

8.6.3 Appendix C: Convergence of CC frequency in an infinite neutral population

In an infinite population with no selection, the frequency of C in $[0,1]$, represented by $x = P_{CC} + P_{CD}/2$, remains constant since then

$$x' - x = \mathbf{E}(\Delta x) = 0 \tag{8.6.25}$$

owing to (8.6.19) as $N \to +\infty$. Moreover, the change in the frequency of CC from time t to time $t + \Delta t$ is given by

$$P'_{CC} - P_{CC} = \mathbf{E}(\Delta P_{CC}) = \frac{(x - P_{CC})^2 - \rho P_{CC}(1 - 2x + P_{CC})}{1 - (1-\rho)P_{CC}} \tag{8.6.26}$$

owing to (8.6.23) as $N \to +\infty$. After algebraic manipulations, this leads to the recurrence equation

$$P'_{CC} = 2x - 1 + \frac{(1-x)^2}{1 - (1-\rho)P_{CC}} = h(P_{CC}). \tag{8.6.27}$$

From the facts that P_{CC}, $P_{CD} = 2(x - P_{CC})$ and $P_{DD} = 1 - 2x + P_{CC}$ are all in $[0,1]$, we have the constraints

$$\max\{2x-1, 0\} \leq P_{CC} \leq x. \tag{8.6.28}$$

Note that $h(0) = x^2 \geq 0$ and $h(2x-1) \geq 2x-1$, so that $h(\max\{2x-1,0\}) \geq \max\{2x-1,0\}$, while

$$h(x) = x - \frac{\rho x(1-x)}{1-(1-\rho)x} \leq x. \tag{8.6.29}$$

On the other hand, the first and second derivatives of h are given by

$$\frac{dh(P_{CC})}{dP_{CC}} = \frac{(1-x)^2(1-\rho)}{(1-(1-\rho)P_{CC})^2} \geq 0 \tag{8.6.30}$$

and

$$\frac{d^2h(P_{CC})}{dP_{CC}^2} = \frac{2(1-x)^2(1-\rho)^2}{(1-(1-\rho)P_{CC})^3} \geq 0, \tag{8.6.31}$$

respectively. By solving the equation $h(P_{CC}) = P_{CC}$, that is,

$$(1-\rho)P_{CC}^2 - (2x(1-\rho)+\rho)P_{CC} + x^2 = 0, \tag{8.6.32}$$

the only equilibrium point of h in the interval $[\max\{2x-1, 0\}, x]$ is found to be

$$P_{CC}^* = x + \frac{\rho}{2(1-\rho)} - \frac{\sqrt{\rho^2 + 4x(1-x)\rho(1-\rho)}}{2(1-\rho)}. \tag{8.6.33}$$

Owing to the above properties, this is a globally stable equilibrium point. At this equilibrium, we get from (8.6.27) that

$$\frac{(1-x)^2}{1-(1-\rho)P_{CC}^*} = P_{CC}^* - 2x + 1, \tag{8.6.34}$$

which simplifies the expressions for $\mathbf{E}(q_{CC})$, $\mathbf{E}(q_{CD})$ and $\mathbf{E}(q_{DD})$ in (8.6.11).

Actually, P_{CC}^* is a uniformly globally stable equilibrium point. As a matter of fact, applying the mean value theorem, there exists $\tilde{P}_{CC}$ between P_{CC} and P_{CC}^* such that

$$P_{CC}' - P_{CC}^* = h(P_{CC}) - h(P_{CC}^*) = (P_{CC} - P_{CC}^*)\frac{dh(\tilde{P}_{CC})}{dP_{CC}} \tag{8.6.35}$$

with

$$\frac{dh(\tilde{P}_{CC})}{dP_{CC}} = \frac{(1-x)^2(1-\rho)}{(1-(1-\rho)\tilde{P}_{CC})^2} \leq \frac{(1-x)^2(1-\rho)}{(1-(1-\rho)x)^2} \leq 1-\rho < 1. \tag{8.6.36}$$

Iterating (8.6.35), we have

$$|P_{CC}^{(n)} - P_{CC}^*| \leq |P_{CC} - P_{CC}^*|(1-\rho)^n \leq (1-\rho)^n \tag{8.6.37}$$

for all integers $n \geq 1$, with $(1-\rho)^n \to 0$ as $n \to \infty$.

8.6.4 Appendix D: Bounds of $f(x, \rho)$

For the additive PD game, the drift function $m(x)$ is in the form

$$m(x) = \sigma\left[x(1-x)(b-c) - \frac{bf(x,\rho)}{2(1-\rho)}\right], \tag{8.6.38}$$

where

$$f(x,\rho) = \sqrt{\rho^2 + 4x(1-x)\rho(1-\rho)} - \rho = \phi(4x(1-x),\rho)(\sqrt{\rho} - \rho) \tag{8.6.39}$$

with

$$\phi(u,\rho) = \frac{\sqrt{\rho + u(1-\rho)} - \sqrt{\rho}}{1 - \sqrt{\rho}} \tag{8.6.40}$$

for $u, \rho \in [0,1]$. We have $\phi(u,0) = \sqrt{u}$ and $\phi(u,1) = \lim_{\rho \to 1} \phi(u,\rho) = u$ by applying L'Hôpital's rule. Moreover,

$$\frac{\partial \phi(u,\rho)}{\partial \rho} = \frac{(1-u)\sqrt{\rho} + u - \sqrt{\rho + u(1-\rho)}}{2(1-\sqrt{\rho})^2 \sqrt{\rho}\sqrt{\rho + u(1-\rho)}} \leq 0, \tag{8.6.41}$$

since

$$(1-u)\sqrt{\rho} + u \leq \sqrt{(1-u)\rho + u} = \sqrt{\rho + u(1-\rho)} \tag{8.6.42}$$

by Jensen's inequality for the concave square root function on $[0,1]$. Therefore, $\phi(u,\rho)$ is a decreasing function of ρ from $\sqrt{u}$ at $\rho = 0$ to u at $\rho = 1$ for every $u \in [0,1]$.

We conclude that

$$\sqrt{4x(1-x)}(\sqrt{\rho} - \rho) \geq f(x,\rho) = \phi(4x(1-x),\rho)(\sqrt{\rho} - \rho) \geq 4x(1-x)(\sqrt{\rho} - \rho), \tag{8.6.43}$$

the upper bound being the limit of $f(x,\rho)$ as $\rho \to 0$ and the lower bound the limit as $\rho \to 1$ (see Figure 8.2).

Acknowledgments

This research was supported in part by NSERC of Canada (Grant no. 8833) and Chinese Academy of Sciences President's International Fellowship Initiative (Grant no. 2016VBA039).

Bibliography

[1] C.A. Aktipis, Know when to walk away: contingent movement and the evolution of cooperation. *Journal of Theoretical Biology* **231**, 249-260 (2004).

[2] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[3] R. Axelrod, W.D. Hamilton, The evolution of cooperation. *Science* **211**, 1390-1396 (1981).

[4] R. Bergmüller, R. Schürch, I.M. Hamilton, Evolutionary causes and consequences of consistent individual variation in cooperative behaviour. *Philosophical Transactions of the Royal Society B* **365**, 2751-2764 (2010)

[5] J. B. Brask, D.P. Croft, M. Edenbrow, R. James, B.H. Bleakley, I.W. Ramnarine, R.J.P. Heathcote, C.R. Tyler, P.B. Hamilton, T. Dabelsteen and S.K. Darden. Evolution of non-kin cooperation: social assortment by cooperative phenotype in guppies. *Royal Society Open Science* **6**: 181493. http://dx.doi.org/10.1098/rsos.181493 (2019).

[6] D.P. Croft, M. Edenbrow, S.K. Darden. Assortment in social networks and the evolution of cooperation. In *Animal Social Networks*, pp. 13-23, J. Krause, R. James, D.W. Franks, D.P. Croft, Eds. (Oxford University Press, Oxford, UK, 2015).

[7] I. Eshel, L.L. Cavalli-Sforza. Assortment of encounters and evolution of cooperativeness. *Proceedings of the National Academy of Sciences USA* **79**, 1331-1335 (1982).

[8] S.N. Ethier, T. Nagylaki. Diffusion approximations of Markov chains with two time scales and applications to population genetics. *Advances in Applied Probability* **12**, 14-49 (1980).

[9] W.J. Ewens, *Mathematical Population Genetics: I Theoretical Introduction* (Springer, New York, 2004).

[10] T. Fujiwara-Greve, M. Okuno-Fujiwara, Voluntarily separable repeated prisoner's dilemma. *The Review of Economic Studies* **76**, 993-1021 (2009).

[11] N. Hayashi, From TIT-for-TAT to OUT-for-TAT. *Sociological Theory and Methods* **8**, 19-32 (1993).

[12] J. Hofbauer, K. Sigmund, *The Theory of Evolution and Dynamical Systems* (Cambridge University Press, Cambridge, 1998).

[13] S.S. Izquierdo, L.R. Izquierdo, F. Vega-Redondo, The option to leave: Conditional dissociation in the evolution of cooperation. *Journal of Theoretical Biology* **267**, 76-84 (2010).

[14] S.S. Izquierdo, L.R. Izquierdo, F. Vega-Redondo, Leave and let leave: A sufficient condition to explain the evolutionary emergence of cooperation. *Journal of Economic Dynamics and Control* **46**, 91-113 (2014).

[15] M. Kimura, Diffusion models in population genetics. *Journal of Applied Probability* **1**(2), 177-232 (1964).

[16] J. Krause, R. James, D.W. Franks, D.P. Croft, *Animal Social Networks* (Oxford University Press, Oxford, UK, 2015).

[17] S. Kurokawa, Three-player repeated games with an opt-out option. *Journal of Theoretical Biology* **480**, 13-22 (2019).

[18] S. Lessard, Evolutionary dynamics in frequency-dependent two-phenotype models. *Theoretical Population Biology* **25**, 210-234 (1984).

[19] C. Li, S. Lessard, Randomized matrix games in a finite population: effect of stochastic fluctuations in the payoffs on the evolution of cooperation. *Theoretical Population Biology* **134**, 77-91 (2020).

[20] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).

[21] M.A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).

[22] M.A. Nowak, R. Highfield, *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed* (Free Press, New York, 2011).

[23] M.A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646-650 (2004).

[24] M.A. Nowak, K. Sigmund, "How populations cohere: Five rules for cooperation" in *Theoretical Ecology - Principles and Applications*, pp. 7-16, R.M. May, A. McLean, Eds. (Oxford University Press, Oxford, 2007).

[25] H. Ohtsuki, C. Hauert, E. Lieberman, M. A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502-505 (2006).

[26] W. Poundstone, *Prisoner's Dilemma* (Doubleday, New York, 1992).

[27] H. Risken, *The Fokker-Planck Equation, Methods of Solution and Applications* (Springer, Berlin, 1992).

[28] R. Schuessler, Exit threats and cooperation under anonymity. *Journal of Conflict Resolution* **33**, 728-749 (1989).

[29] R.L. Trivers, Evolution of reciprocal altruism. *Quarterly Review of Biology* **46**, 35-57 (1971).

[30] B.-Y. Zhang, S.-J. Fan, C. Li, X.-D. Zheng, J.-Z. Bao, R. Cressman, Y. Tao, Opting out against defection leads to stable coexistence with cooperation. *Scientific Reports* **6**, 35902 (2016).

[31] X.-D. Zheng, C. Li, J.-R. Yu, S.-C, Wang, S.-J. Fan, B.-Y. Zhang, Y. Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. *Journal of Theoretical Biology* **420**, 12-17 (2017).

Chapter 9

Conclusion

9.1 Part 1

In part 1 of the thesis we have studied matrix games in a stochastic environment. For general 2×2 randomized matrix games in infinite populations undergoing discrete non-overlapping generations, Articles 1, 2, and 3 have established conditions for boundary and interior equilibria to be SLS or SLU. In Article 1, the expected payoff is used directly as fitness, that is, $fitness = payoff$. This models strong selection. In Article 2, an exponential function of the payoff is used as fitness, that is, $fitness = \text{Exp}(payoff)$. This nonlinear function is approached by the linear function $1 + payoff$ in the case of weak selection. In Article 3, it is the function $fitness = (1-w)+w \times payoff$, where the intensity of selection w is small, which is used as fitness. In Articles 4 and 5, we have analysed randomized Prisoner's Dilemma games in an infinite and a finite population, respectively, and revealed that increasing the variance in the payoffs for defection is conducive to the evolution of cooperation.

In a 2×2 randomized matrix game, the conditions for a constant interior equilibrium to exist and be SLS depend on the structure of the payoff matrix, since we must have $(d_t - b_t)/(a_t - c_t) = \hat{u} > 0$ for all $t \geq 0$. Otherwise, the population state may wander around a locally stable interior equilibrium with respect to the mean payoff matrix without ever converging to it. In this case, numerical simulations show that the population state tends to stay in a region surrounding the interior equilibrium whose size depends on the variances of the payoffs. It would be interesting to study this phenomenon in more detail.

In Articles 1 and 2, we have noticed that it is possible for both fixation states and a constant interior equilibrium in a 2×2 randomized matrix game to be simultaneously SLS with no other interior constant equilibrium in the sense that the expected payoffs of the two strategies are always equal. Such a situation never occurs in the deterministic dynamics with constant payoffs, where at least one unstable equilibrium must exist between two stable equilibria. However, there might be some unstable region between two SLS equilibria in the case of random payoffs whose characteristics remain to be defined and studied.

In the case of a randomized Prisoner's Dilemma (RPD), we have shown that the evolution of cooperation is promoted when the variability in the payoffs for defection is greater than that for cooperation. There might be several explanations for this result. It seems that the effects of increases and decreases of the payoffs are asymmetric and detrimental with respect to evolution. Therefore, decreasing the variance in payoffs for cooperation or increasing the variance in payoffs for defection should be beneficial for the evolution of cooperation. It remains to find conditions that explain why defectors have greater uncertainty in their payoffs. Besides environmental noise, uncertainty in payoffs may be introduced in different ways. It may be due to the effects of interacting rules, such as the opting-out strategy studied in part 2 where defectors may face more uncertainty than cooperators, or punishment strategies from which defectors might suffer. These effects remain to be studied.

Note that all the articles above studied only 2×2 matrix games. It would be worth extending these studies to more general games, that is, multi-strategy games such as the Rock-Paper-Scissors game, multi-player games such as Public Goods games, and asymmetric games such as the Ultimatum game (M.A. Nowak, K.M. Page, K. Sigmund Fairness versus reason in the ultimatum game. Science. 289(5485):1773-1775 (2000)).

9.2 Part 2

In part 2 of the thesis, we have studied the effects of opting-out in a repeated Prisoner's Dilemma game with additive cost c and benefit b for cooperation in both infinite and finite populations. Under the assumption that every individual adopts the opting-out strategy, that is, wants to continue the interaction as long as the opponent cooperates, but stop it as soon as the opponent defects, we have shown that the evolution of cooperation is promoted if the benefit-to-cost ratio $r = b/c$ is large enough or the probability ρ that an interaction between two players willing to continue breaks out is small enough. More specifically, if $r > (1 + \sqrt{\rho})/(1 - \sqrt{\rho})$, then a stable interior equilibrium exists in the case of an infinite population and selection favors the evolution of cooperation in the case of a finite population.

Consider, for instance, the two strategies tit-for-tat (TFT) and always-defect (AllD) in a repeated PD game. The TFT strategy is an ESS if $r > 1/(1 - \rho)$; selection favors the evolution of TFT more than the evolution of AllD if $r > (1 + \rho)/(1 - \rho)$; the evolution of TFT is fully favored by selection if $r > (1 + 2\rho)/(1 - \rho)$. Since we have the inequality $(1 + \sqrt{\rho})/(1 - \sqrt{\rho}) > (1 + 2\rho)/(1 - \rho)$ for $0 < \rho < 1$, we conclude that the condition for selection to favor the evolution of cooperation is more stringent with opting-out than with TFT. This is somehow in agreement with experimental results reported in Article 7. However, this does not necessarily mean that TFT is better than the opting-out strategy for the evolution of cooperation. The Monte Carlo simulations in Article 6 (see also Izquierdo *et al.*, 2010) have shown that the opting-out strategy can prevail more often than TFT in the population when both strategies are available in strategy-pool. However, there is no theoretical analysis to support this result yet. Further studies that involve various strategies, such as TFT and PAVLOV, together with opting-out are worth considering in the future.

The opting-out strategy provides the opportunity not only for cooperators to find cooperative partners but also for defectors who have an even greater advantage to do so. Especially, when the frequency of cooperators is large, defectors can easily invade the population by frequently changing partner and taking advantage of it in the first round. In the real world, however, ending an interaction with someone might incur a cost since there is a risk of not finding a new partner for the next round, and this is not considered in our model. Adding a cost function would be a way to improve it. Moreover, the cost could be different for cooperators and defectors, which could favor the evolution of cooperation.

Moreover, our experiment results indicate that the probability of interrupting an interaction when the opponent defects is significantly different from our assumption in our theoretical model which is based on a rational choice. This probability may even be different for a cooperator C and a defector D. A reasonable improvement of our model would be to consider different probabilities ρ_{CC}, ρ_{CD} and ρ_{DD}, where ρ_{XY} represents the probability for an individual adopting an action X to interrupt an interaction when facing an individual adopting an action Y. These parameters could themselves be subject to evolution.

Our articles reveal that the long-term coexistence of cooperation and defection in a population is possible when introducing the opting-out strategy. This result help us to understand the reason that why it is hard to observe a society with full cooperation. In the case if almost every one cooperates in the population, defector can easily find a cooperator and exploit in the first round of a game. Thus defector is benefited in this situation and won't extinct. Finally, we have focussed on PD games with two players. There are other multi-player games to study the evolution of cooperation, such as Public Goods games, and it would be interesting to study the effects of opting-out in such games.

Declarations and permissions

This section contains the declarations and permissions from the co-authors of the articles and the editors or publisher of the journals where they were published or submitted.

Declarations of the co-authors:

(1) Sabin Lessard for Articles 1, 2, 3, 4, 5, 8.

(2) Yi Tao for Articles 1, 2, 3, 4, 6, 7.

(3) Xiu-Deng Zheng for Articles 1, 2, 3, 4, 6, 7.

(4) Tian-Jiao Feng for Article 3.

(5) Ming-Yang Wang for Article 3.

(6) Ting Ji for Article 4.

(7) Qiao-Qiao He for Article 4.

(8) Jie-Ru Yu for Article 6.

(9) Shi-Chang Wang for Article 6.

(10) Song-Jia Fan for Articles 6, 7.

(11) Bo-Yu Zhang for Articles 6, 7.

(12) Jian-Zhang Bao for Article 7.

(13) Ross Cressman for Article 7.

Permissions from the editors or publisher:

(14) Michael Thoennessen (Editor in Chief, American Physical Society) for Articles 1, 2, 3.

(15) Kaveri Thakuria (Permissions Helpdesk of Elsevier) for Article 4, 5, 6, 8.

(16) Richard White (Chief Editor, Scientific Reports) for Article 7.

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

1. Identification of the coauthor (complete name and location)

   Sabin Lessard (Professor)
   Department of Mathematics and Statistics
   University of Montreal
   Montreal, Quebec H3C 3J7
   Canada

2. Identification of the articles

   Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Evolutionary stability concepts in a stochastic environment. *Physical Review E* **96**, 032414 (2017).

   Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Environmental Noise Could Promote Stochastic Local Stability of Behavioral Diversity Evolution. *Physical Review Letters* **120**, 218101 (2018).

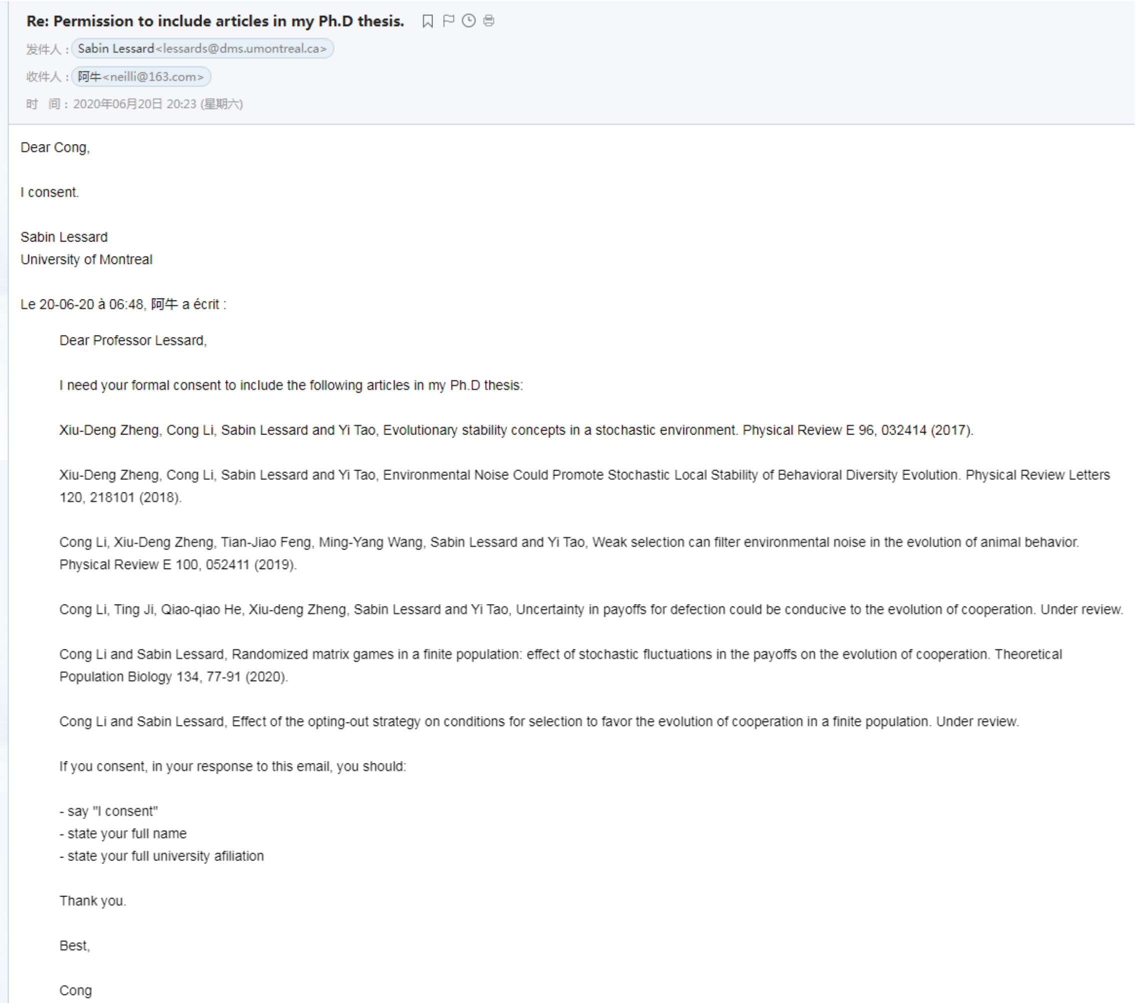
   Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. *Physical Review E* **100**, 052411 (2019).

   Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review.

   Cong Li and Sabin Lessard, Randomized matrix games in a finite population: effect of stochastic fluctuations in the payoffs on the evolution of cooperation. *Theoretical Population Biology* **134**, 77-91 (2020).

   Cong Li and Sabin Lessard, Effect of the opting-out strategy on conditions for selection to favor the evolution of cooperation in a finite population. Under review.

3. Declaration



Re: Permission to include articles in my Ph.D thesis.

发件人：Sabin Lessard<lessards@dms.umontreal.ca>
收件人：阿牛<neilli@163.com>
时　间：2020年06月20日 20:23 (星期六)

Dear Cong,

I consent.

Sabin Lessard
University of Montreal

Le 20-06-20 à 06:48, 阿牛 a écrit :

> Dear Professor Lessard,
>
> I need your formal consent to include the following articles in my Ph.D thesis:
>
> Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Evolutionary stability concepts in a stochastic environment. Physical Review E 96, 032414 (2017).
>
> Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Environmental Noise Could Promote Stochastic Local Stability of Behavioral Diversity Evolution. Physical Review Letters 120, 218101 (2018).
>
> Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. Physical Review E 100, 052411 (2019).
>
> Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review.
>
> Cong Li and Sabin Lessard, Randomized matrix games in a finite population: effect of stochastic fluctuations in the payoffs on the evolution of cooperation. Theoretical Population Biology 134, 77-91 (2020).
>
> Cong Li and Sabin Lessard, Effect of the opting-out strategy on conditions for selection to favor the evolution of cooperation in a finite population. Under review.
>
> If you consent, in your response to this email, you should:
>
> - say "I consent"
> - state your full name
> - state your full university afiliation
>
> Thank you.
>
> Best,
>
> Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

1. Identification of the coauthor (complete name and location)

   Yi Tao (Professor)
   Key Lab of Animal Ecology and Conservation Biology
   Institute of Zoology, Chinese Academy of Science
   Beijing, P. R. China

2. Identification of the articles

   Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Evolutionary stability concepts in a stochastic environment. *Physical Review E* **96**, 032414 (2017).

   Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Environmental Noise Could Promote Stochastic Local Stability of Behavioral Diversity Evolution. *Physical Review Letters* **120**, 218101 (2018).

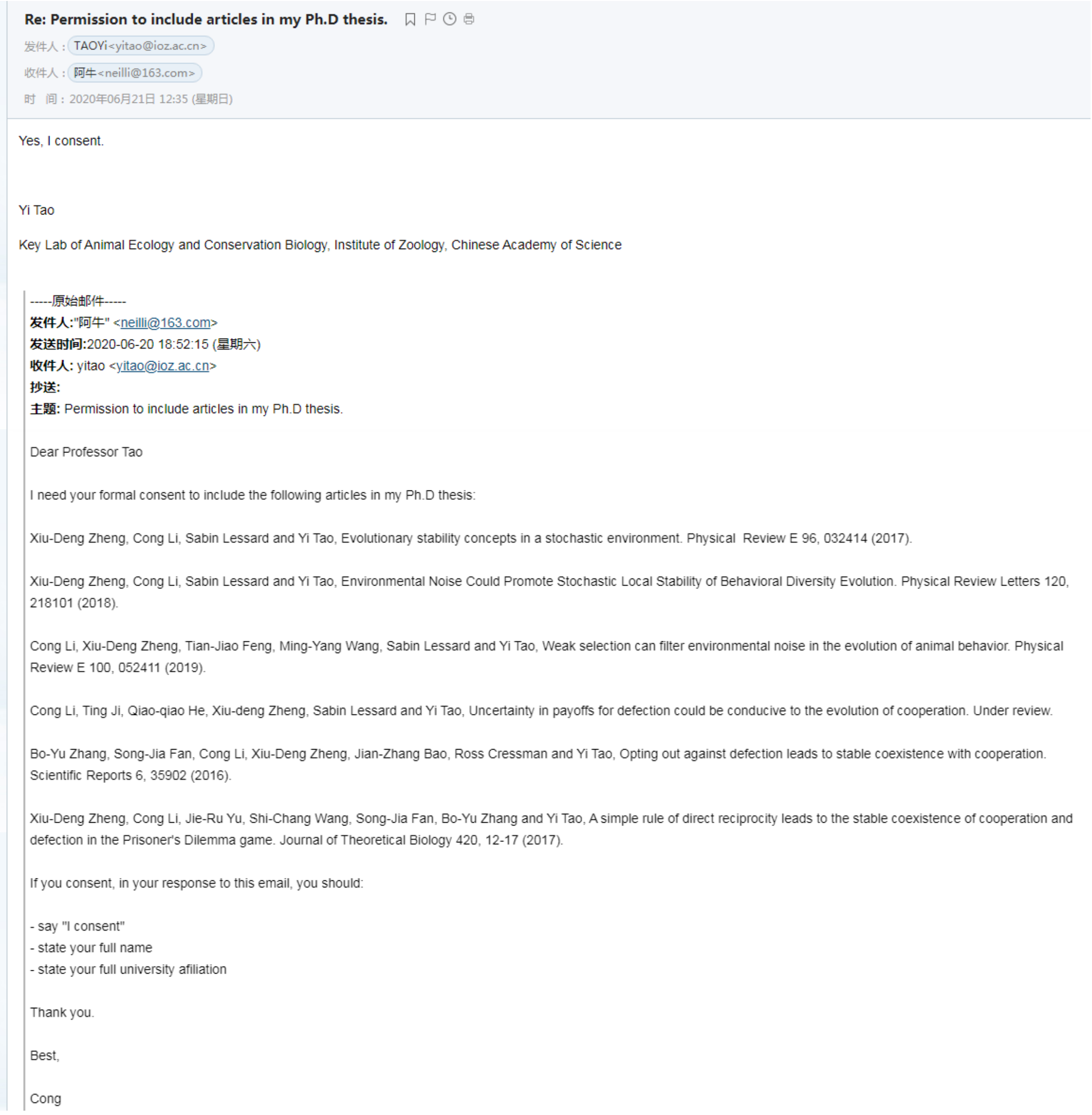
   Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. *Physical Review E* **100**, 052411 (2019).

   Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review.

   Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. *Journal of Theoretical Biology* **420**, 12-17 (2017).

   Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. *Scientific Reports* **6**, 35902 (2016).

3. Declaration



Re: Permission to include articles in my Ph.D thesis.

发件人：TAOYi<yitao@ioz.ac.cn>

收件人：阿牛<neilli@163.com>

时　间：2020年06月21日 12:35 (星期日)

Yes, I consent.

Yi Tao

Key Lab of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Science

> -----原始邮件-----
> **发件人:**"阿牛" <neilli@163.com>
> **发送时间:**2020-06-20 18:52:15 (星期六)
> **收件人:** yitao <yitao@ioz.ac.cn>
> **抄送:**
> **主题:** Permission to include articles in my Ph.D thesis.
>
> Dear Professor Tao
>
> I need your formal consent to include the following articles in my Ph.D thesis:
>
> Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Evolutionary stability concepts in a stochastic environment. Physical Review E 96, 032414 (2017).
>
> Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Environmental Noise Could Promote Stochastic Local Stability of Behavioral Diversity Evolution. Physical Review Letters 120, 218101 (2018).
>
> Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. Physical Review E 100, 052411 (2019).
>
> Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review.
>
> Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. Scientific Reports 6, 35902 (2016).
>
> Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. Journal of Theoretical Biology 420, 12-17 (2017).
>
> If you consent, in your response to this email, you should:
>
> - say "I consent"
> - state your full name
> - state your full university affiliation
>
> Thank you.
>
> Best,
>
> Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

1. Identification of the coauthor (complete name and location)

   Xiu-Deng Zheng (Associate Professor)
   Key Lab of Animal Ecology and Conservation Biology
   Institute of Zoology, Chinese Academy of Science
   Beijing, P. R. China

2. Identification of the articles

   Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Evolutionary stability concepts in a stochastic environment. *Physical Review E* **96**, 032414 (2017).

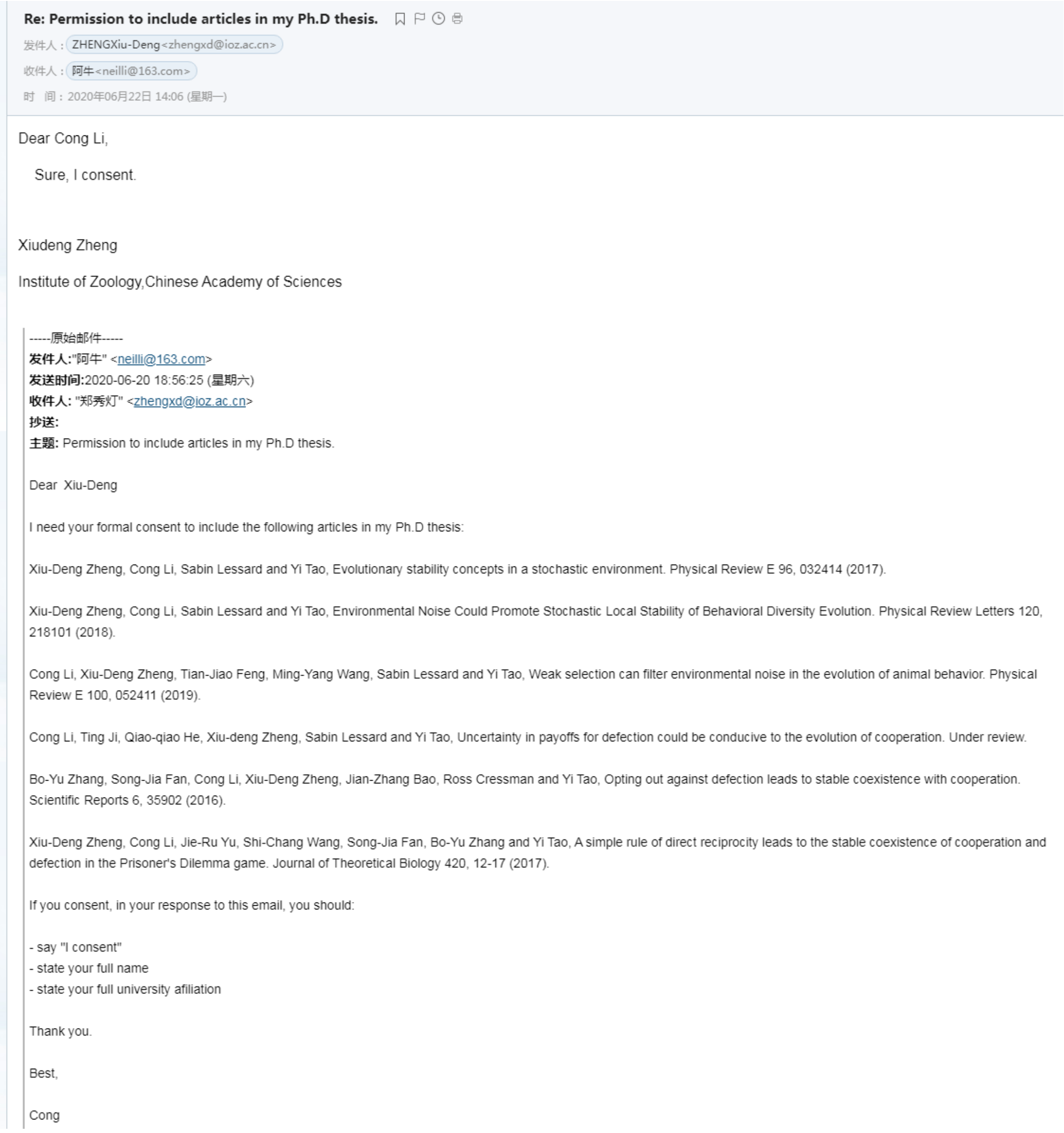
   Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Environmental Noise Could Promote Stochastic Local Stability of Behavioral Diversity Evolution. *Physical Review Letters* **120**, 218101 (2018).

   Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. *Physical Review E* **100**, 052411 (2019).

   Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review

   Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. *Journal of Theoretical Biology* **420**, 12-17 (2017).

   Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. *Scientific Reports* **6**, 35902 (2016).



Re: Permission to include articles in my Ph.D thesis.

发件人：ZHENGXiu-Deng<zhengxd@ioz.ac.cn>

收件人：阿牛<neilli@163.com>

时　间：2020年06月22日 14:06 (星期一)

Dear Cong Li,

Sure, I consent.

Xiudeng Zheng

Institute of Zoology,Chinese Academy of Sciences

-----原始邮件-----
发件人:"阿牛" <neilli@163.com>
发送时间:2020-06-20 18:56:25 (星期六)
收件人: "郑秀灯" <zhengxd@ioz.ac.cn>
抄送:
主题: Permission to include articles in my Ph.D thesis.

Dear Xiu-Deng

I need your formal consent to include the following articles in my Ph.D thesis:

Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Evolutionary stability concepts in a stochastic environment. Physical Review E 96, 032414 (2017).

Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Environmental Noise Could Promote Stochastic Local Stability of Behavioral Diversity Evolution. Physical Review Letters 120, 218101 (2018).

Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. Physical Review E 100, 052411 (2019).

Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review.

Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. Scientific Reports 6, 35902 (2016).

Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. Journal of Theoretical Biology 420, 12-17 (2017).

If you consent, in your response to this email, you should:

- say "I consent"
- state your full name
- state your full university afiliation

Thank you.

Best,

Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

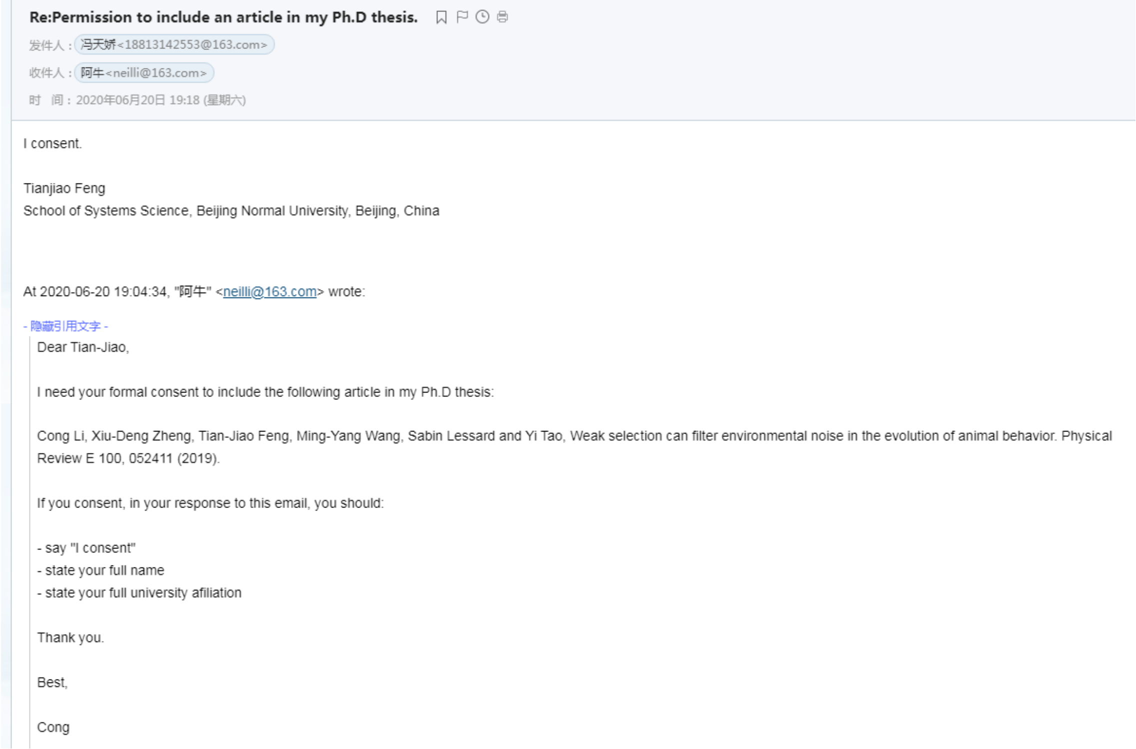
1. Identification of the coauthor (complete name and location)

   Tian-Jiao Feng (Master)
   School of Systems Science
   Beijing Normal University
   Beijing, P. R. China

2. Identification of the articles

   Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. *Physical Review E* **100**, 052411 (2019).

3. Declaration



Re:Permission to include an article in my Ph.D thesis.

发件人：冯天娇<18813142553@163.com>
收件人：阿牛<neilli@163.com>
时　间：2020年06月20日 19:18 (星期六)

I consent.

Tianjiao Feng
School of Systems Science, Beijing Normal University, Beijing, China

At 2020-06-20 19:04:34, "阿牛" <neilli@163.com> wrote:

- 隐藏引用文字 -

> Dear Tian-Jiao,
>
> I need your formal consent to include the following article in my Ph.D thesis:
>
> Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. Physical Review E 100, 052411 (2019).
>
> If you consent, in your response to this email, you should:
>
> - say "I consent"
> - state your full name
> - state your full university afiliation
>
> Thank you.
>
> Best,
>
> Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

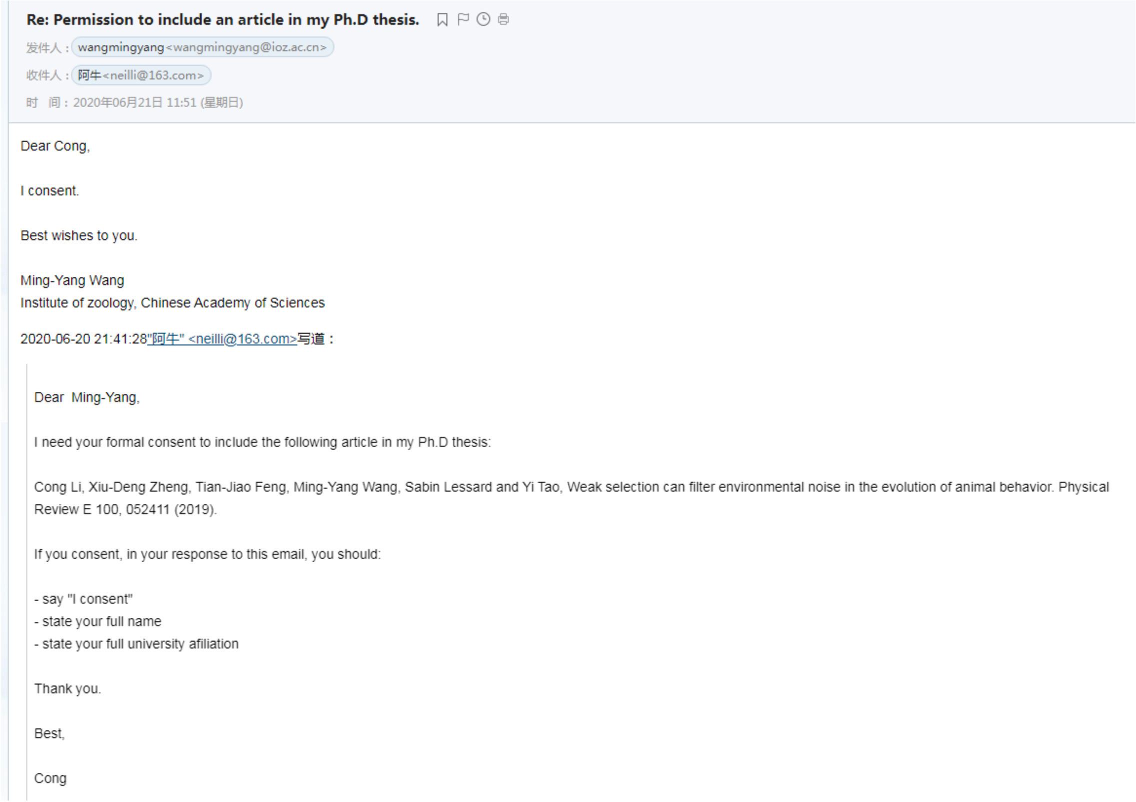
1. Identification of the coauthor (complete name and location)

   Ming-Yang Wang (Master)
   Key Lab of Animal Ecology and Conservation Biology
   Institute of Zoology, Chinese Academy of Science
   Beijing, P. R. China

2. Identification of the articles

   Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. *Physical Review E* **100**, 052411 (2019).

3. Declaration



Re: Permission to include an article in my Ph.D thesis.

发件人：wangmingyang<wangmingyang@ioz.ac.cn>
收件人：阿牛<neilli@163.com>
时　间：2020年06月21日 11:51 (星期日)

Dear Cong,

I consent.

Best wishes to you.

Ming-Yang Wang
Institute of zoology, Chinese Academy of Sciences

2020-06-20 21:41:28"阿牛" <neilli@163.com>写道：

> Dear Ming-Yang,
>
> I need your formal consent to include the following article in my Ph.D thesis:
>
> Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. Physical Review E 100, 052411 (2019).
>
> If you consent, in your response to this email, you should:
>
> - say "I consent"
> - state your full name
> - state your full university afiliation
>
> Thank you.
>
> Best,
>
> Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

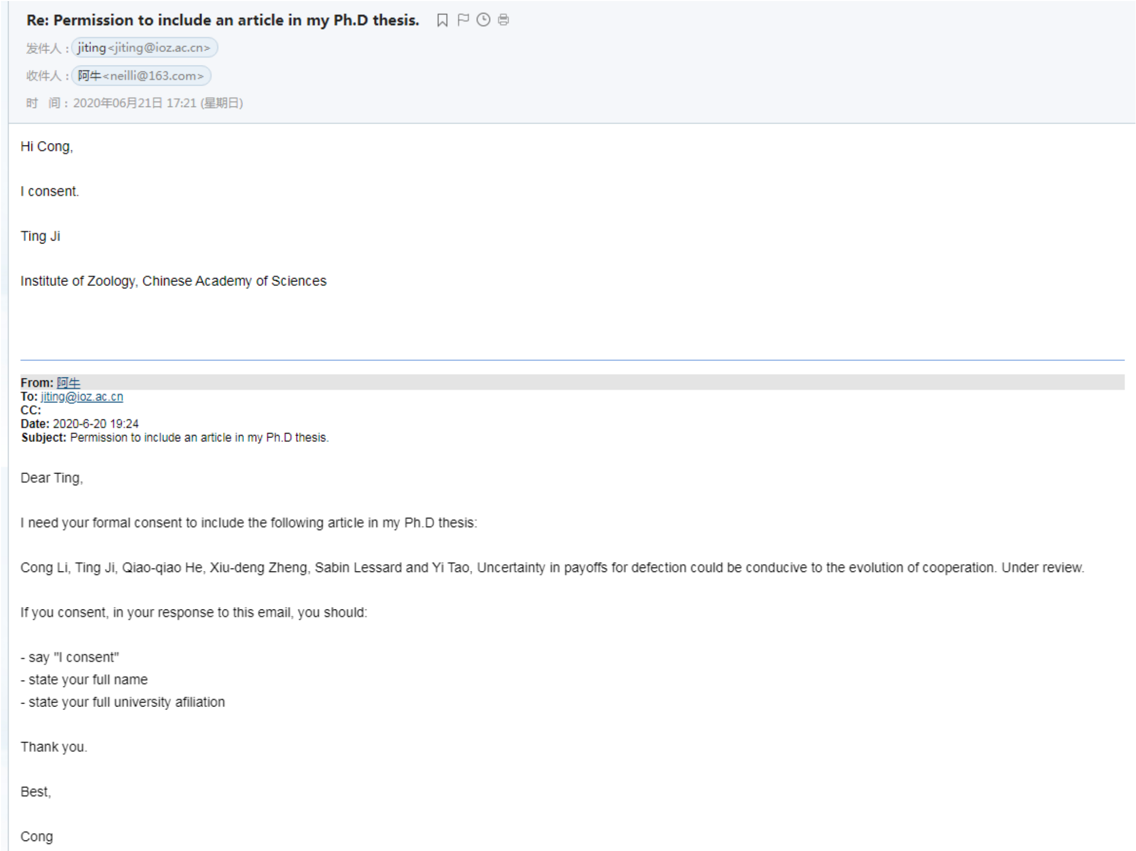
1. Identification of the coauthor (complete name and location)

   Ting Ji (Associate Professor)
   Key Lab of Animal Ecology and Conservation Biology
   Institute of Zoology, Chinese Academy of Science
   Beijing, P. R. China

2. Identification of the articles

   Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review.

3. Declaration



Re: Permission to include an article in my Ph.D thesis.

发件人：jiting<jiting@ioz.ac.cn>
收件人：阿牛<neilli@163.com>
时　间：2020年06月21日 17:21 (星期日)

Hi Cong,

I consent.

Ting Ji

Institute of Zoology, Chinese Academy of Sciences

From: 阿牛
To: jiting@ioz.ac.cn
CC:
Date: 2020-6-20 19:24
Subject: Permission to include an article in my Ph.D thesis.

Dear Ting,

I need your formal consent to include the following article in my Ph.D thesis:

Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review.

If you consent, in your response to this email, you should:

- say "I consent"
- state your full name
- state your full university afiliation

Thank you.

Best,

Cong

DECLARATION FROM A COAUTHOR FOR
THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

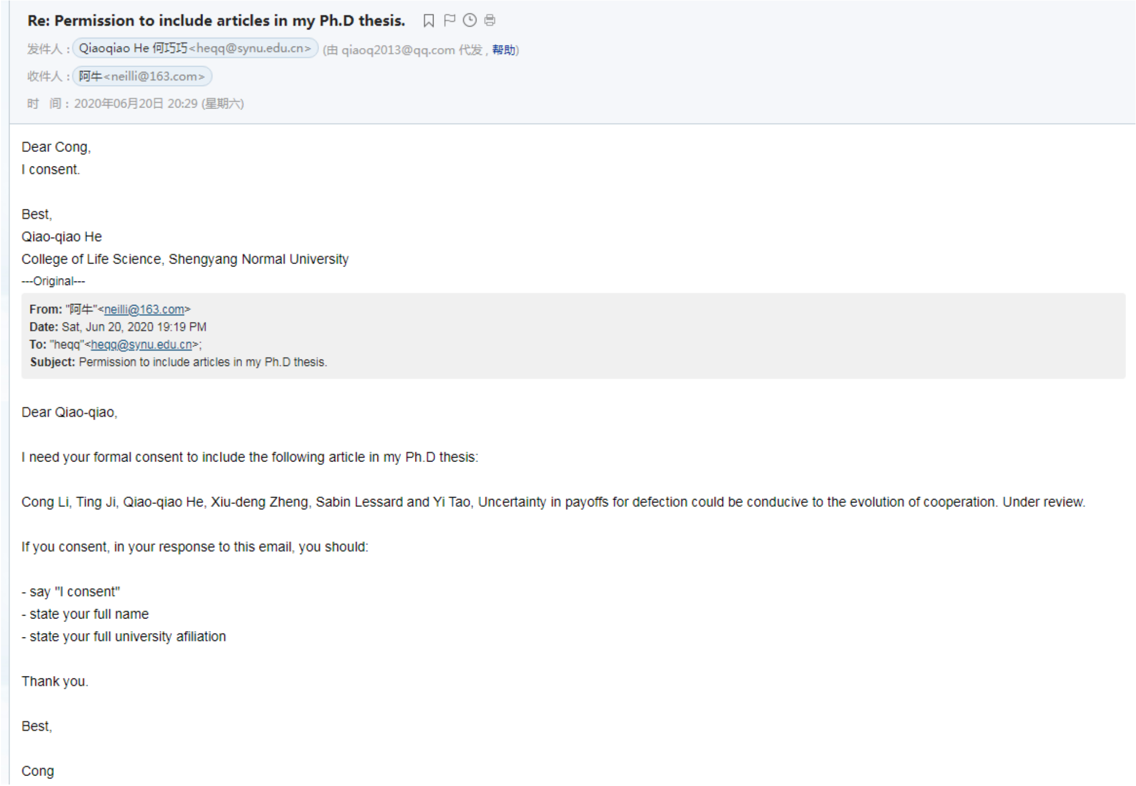
1. Identification of the coauthor (complete name and location)

   Qiao-qiao He (Associate Professor)
   College of Life Science
   Shenyang Normal University
   Shenyang, Liaoning
   P. R. China

2. Identification of the articles

   Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review.

3. Declaration



Re: Permission to include articles in my Ph.D thesis.

发件人：Qiaoqiao He 何巧巧<heqq@synu.edu.cn> (由 qiaoq2013@qq.com 代发，帮助)

收件人：阿牛<neilli@163.com>

时　间：2020年06月20日 20:29 (星期六)

Dear Cong,
I consent.

Best,
Qiao-qiao He
College of Life Science, Shengyang Normal University
---Original---

From: "阿牛"<neilli@163.com>
Date: Sat, Jun 20, 2020 19:19 PM
To: "heqq"<heqq@synu.edu.cn>;
Subject: Permission to include articles in my Ph.D thesis.

Dear Qiao-qiao,

I need your formal consent to include the following article in my Ph.D thesis:

Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review.

If you consent, in your response to this email, you should:

- say "I consent"
- state your full name
- state your full university afiliation

Thank you.

Best,

Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

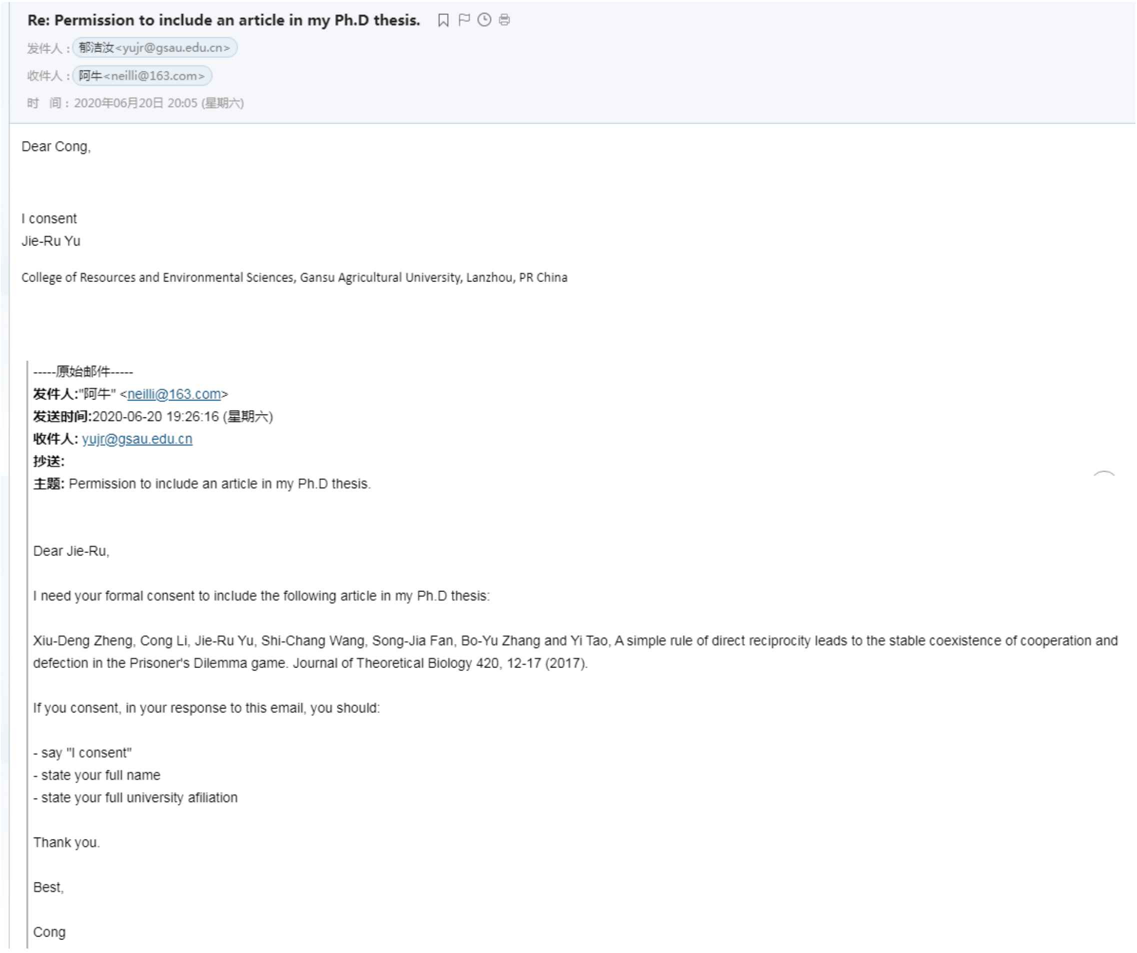
1. Identification of the coauthor (complete name and location)

   Jie-Ru Yu (Associate Professor)
   College of Resources and Environmental Sciences
   Gansu Agricultural University
   Lanzhou, Gansu
   P. R. China

2. Identification of the articles

   Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. *Journal of Theoretical Biology* **420**, 12-17 (2017).

3. Declaration



Re: Permission to include an article in my Ph.D thesis.

发件人：郁洁汝<yujr@gsau.edu.cn>
收件人：阿牛<neilli@163.com>
时　间：2020年06月20日 20:05 (星期六)

Dear Cong,

I consent
Jie-Ru Yu

College of Resources and Environmental Sciences, Gansu Agricultural University, Lanzhou, PR China

> -----原始邮件-----
> **发件人:**"阿牛" <neilli@163.com>
> **发送时间:**2020-06-20 19:26:16 (星期六)
> **收件人:** yujr@gsau.edu.cn
> **抄送:**
> **主题:** Permission to include an article in my Ph.D thesis.
>
> Dear Jie-Ru,
>
> I need your formal consent to include the following article in my Ph.D thesis:
>
> Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. Journal of Theoretical Biology 420, 12-17 (2017).
>
> If you consent, in your response to this email, you should:
>
> - say "I consent"
> - state your full name
> - state your full university afiliation
>
> Thank you.
>
> Best,
>
> Cong

DECLARATION FROM A COAUTHOR FOR
THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

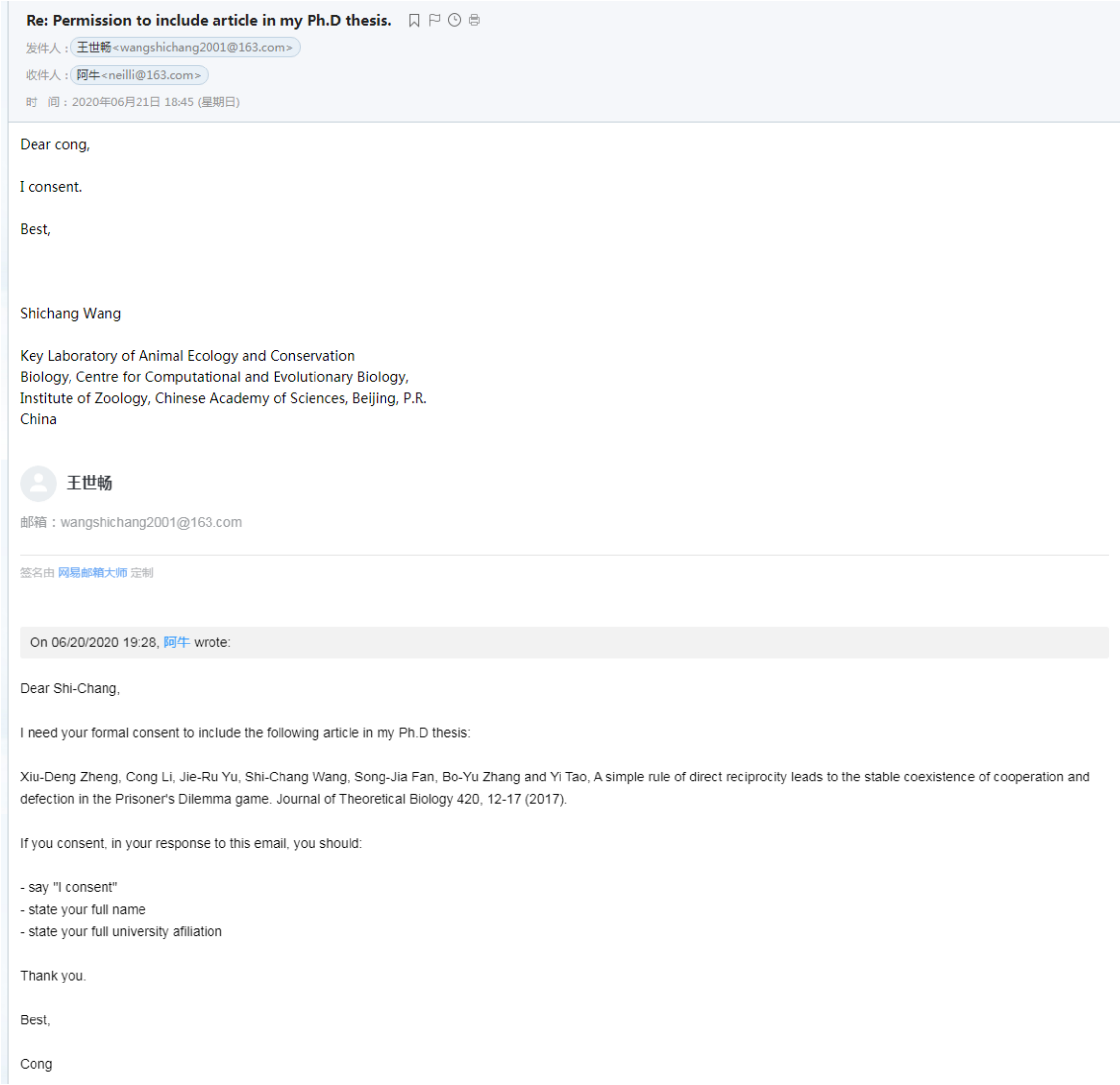
1. Identification of the coauthor (complete name and location)

   Shi-Chang Wang (Ph.D)
   Key Lab of Animal Ecology and Conservation Biology
   Institute of Zoology, Chinese Academy of Science
   Beijing, P. R. China

2. Identification of the articles

   Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. *Journal of Theoretical Biology* **420**, 12-17 (2017).

3. Declaration



Re: Permission to include article in my Ph.D thesis.

发件人：王世畅<wangshichang2001@163.com>
收件人：阿牛<neilli@163.com>
时　间：2020年06月21日 18:45 (星期日)

Dear cong,

I consent.

Best,

Shichang Wang

Key Laboratory of Animal Ecology and Conservation
Biology, Centre for Computational and Evolutionary Biology,
Institute of Zoology, Chinese Academy of Sciences, Beijing, P.R.
China

王世畅

邮箱：wangshichang2001@163.com

签名由 网易邮箱大师 定制

On 06/20/2020 19:28, 阿牛 wrote:

Dear Shi-Chang,

I need your formal consent to include the following article in my Ph.D thesis:

Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. Journal of Theoretical Biology 420, 12-17 (2017).

If you consent, in your response to this email, you should:

- say "I consent"
- state your full name
- state your full university afiliation

Thank you.

Best,

Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

1. Identification of the coauthor (complete name and location)

   Song-Jia Fan (Ph.D)

   Key Lab of Animal Ecology and Conservation Biology

   Institute of Zoology, Chinese Academy of Science

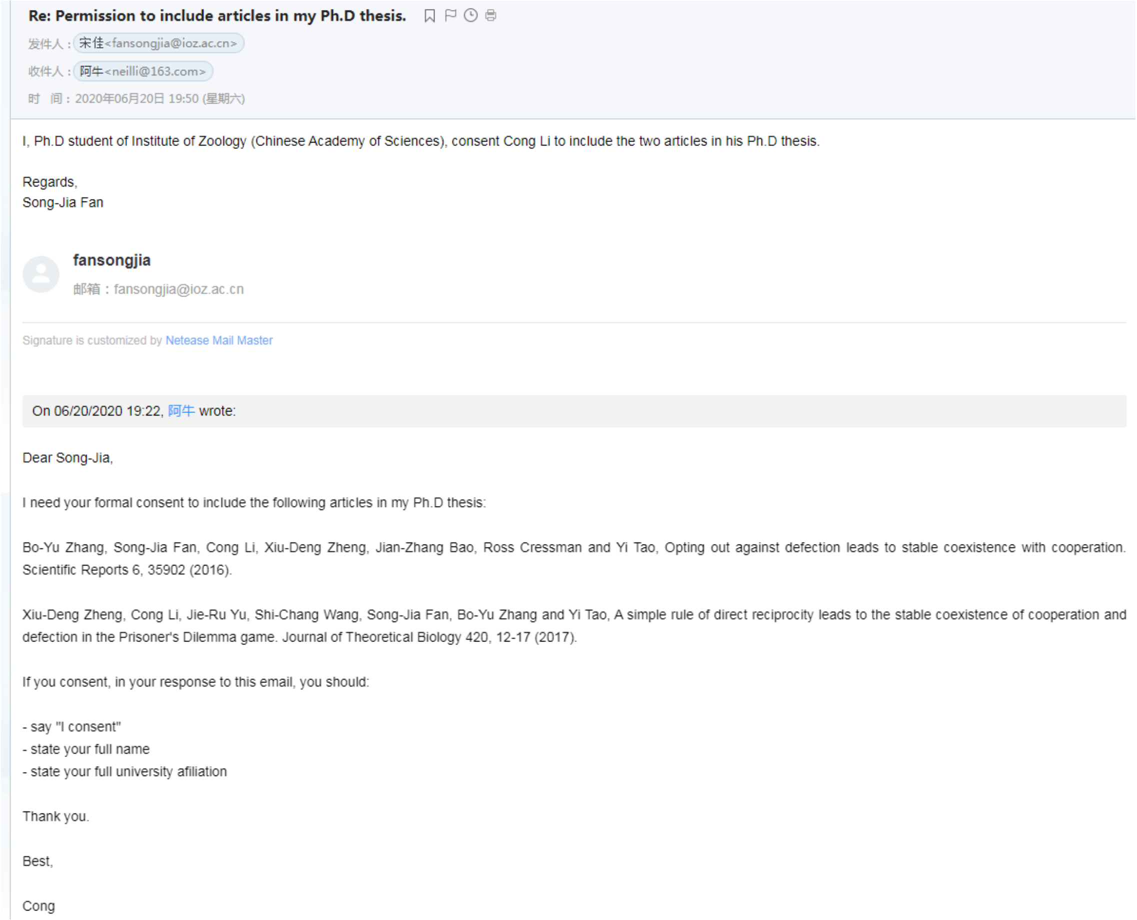
   Beijing, P. R. China

2. Identification of the articles

   Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. *Scientific Reports* **6**, 35902 (2016).

   Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. *Journal of Theoretical Biology* **420**, 12-17 (2017).

3. Declaration



Re: Permission to include articles in my Ph.D thesis.

发件人：宋佳<fansongjia@ioz.ac.cn>
收件人：阿牛<neilli@163.com>
时　间：2020年06月20日 19:50 (星期六)

I, Ph.D student of Institute of Zoology (Chinese Academy of Sciences), consent Cong Li to include the two articles in his Ph.D thesis.

Regards,
Song-Jia Fan

fansongjia
邮箱：fansongjia@ioz.ac.cn

Signature is customized by Netease Mail Master

On 06/20/2020 19:22, 阿牛 wrote:

Dear Song-Jia,

I need your formal consent to include the following articles in my Ph.D thesis:

Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. Scientific Reports 6, 35902 (2016).

Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. Journal of Theoretical Biology 420, 12-17 (2017).

If you consent, in your response to this email, you should:

- say "I consent"
- state your full name
- state your full university afiliation

Thank you.

Best,

Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

1. Identification of the coauthor (complete name and location)

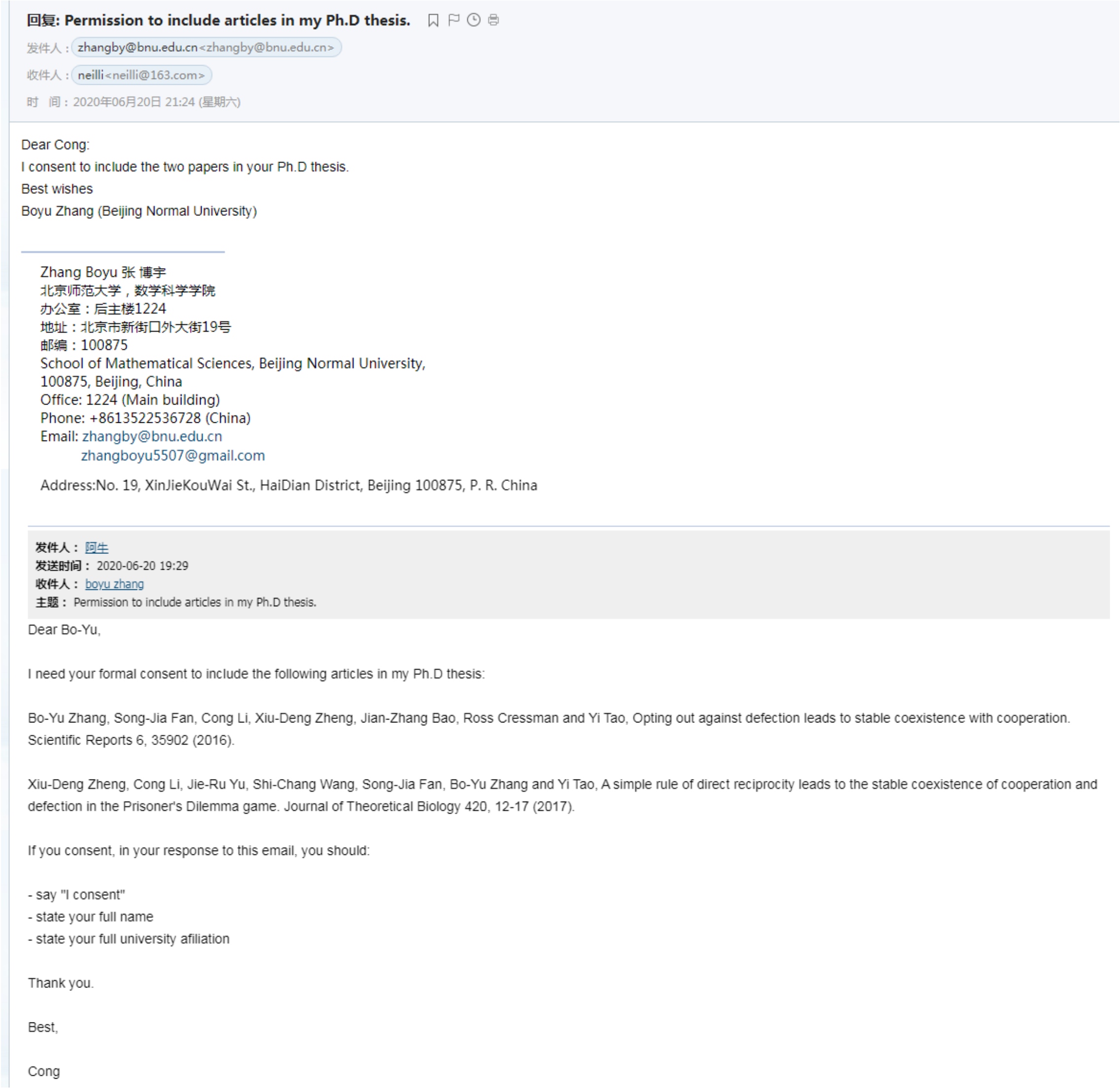
   Bo-Yu Zhang (Professor)
   School of Mathematical Sciences
   Beijing Normal University
   Beijing, P. R. China

2. Identification of the articles

   Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. *Scientific Reports* **6**, 35902 (2016).

   Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. *Journal of Theoretical Biology* **420**, 12-17 (2017).

3. Declaration



回复: Permission to include articles in my Ph.D thesis.

发件人：zhangby@bnu.edu.cn<zhangby@bnu.edu.cn>
收件人：neilli<neilli@163.com>
时　间：2020年06月20日 21:24 (星期六)

Dear Cong:
I consent to include the two papers in your Ph.D thesis.
Best wishes
Boyu Zhang (Beijing Normal University)

Zhang Boyu 张 博宇
北京师范大学，数学科学学院
办公室：后主楼1224
地址：北京市新街口外大街19号
邮编：100875
School of Mathematical Sciences, Beijing Normal University,
100875, Beijing, China
Office: 1224 (Main building)
Phone: +8613522536728 (China)
Email: zhangby@bnu.edu.cn
zhangboyu5507@gmail.com

Address:No. 19, XinJieKouWai St., HaiDian District, Beijing 100875, P. R. China

发件人： 阿生
发送时间： 2020-06-20 19:29
收件人： boyu zhang
主题： Permission to include articles in my Ph.D thesis.

Dear Bo-Yu,

I need your formal consent to include the following articles in my Ph.D thesis:

Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. Scientific Reports 6, 35902 (2016).

Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. Journal of Theoretical Biology 420, 12-17 (2017).

If you consent, in your response to this email, you should:

- say "I consent"
- state your full name
- state your full university afiliation

Thank you.

Best,

Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

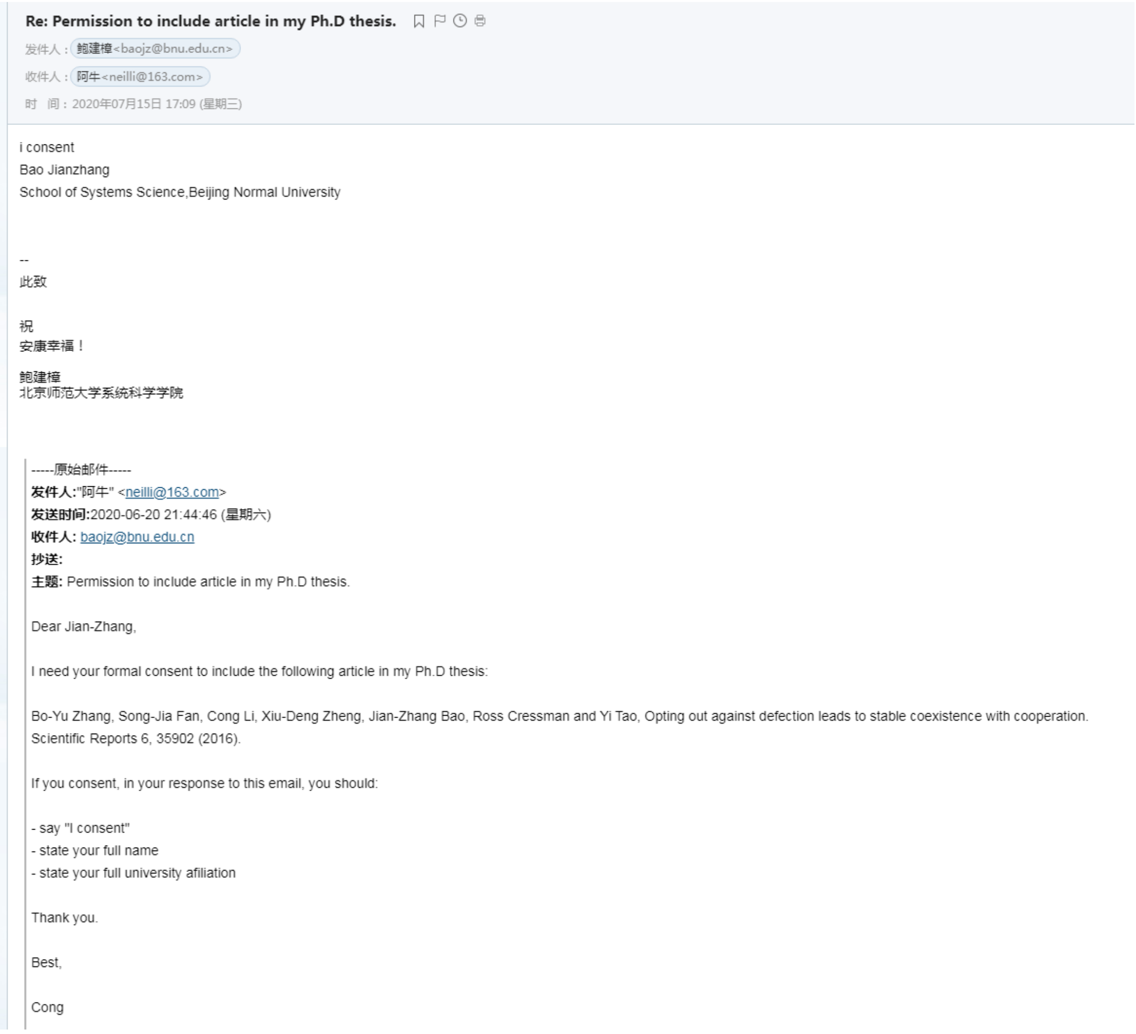
1. Identification of the coauthor (complete name and location)

   Jian-Zhang Bao (Lecturer)
   School of Systems Science
   Beijing Normal University
   Beijing, P. R. China

2. Identification of the articles

   Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. *Scientific Reports* **6**, 35902 (2016).

3. Declaration



Re: Permission to include article in my Ph.D thesis.

发件人：鲍建樟<baojz@bnu.edu.cn>

收件人：阿牛<neilli@163.com>

时　间：2020年07月15日 17:09 (星期三)

i consent
Bao Jianzhang
School of Systems Science,Beijing Normal University

--
此致

祝
安康幸福！

鲍建樟
北京师范大学系统科学学院

> -----原始邮件-----
> **发件人:**"阿牛" <neilli@163.com>
> **发送时间:**2020-06-20 21:44:46 (星期六)
> **收件人:** baojz@bnu.edu.cn
> **抄送:**
> **主题:** Permission to include article in my Ph.D thesis.
>
> Dear Jian-Zhang,
>
> I need your formal consent to include the following article in my Ph.D thesis:
>
> Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. Scientific Reports 6, 35902 (2016).
>
> If you consent, in your response to this email, you should:
>
> - say "I consent"
> - state your full name
> - state your full university afiliation
>
> Thank you.
>
> Best,
>
> Cong

DECLARATION FROM A COAUTHOR FOR THE INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

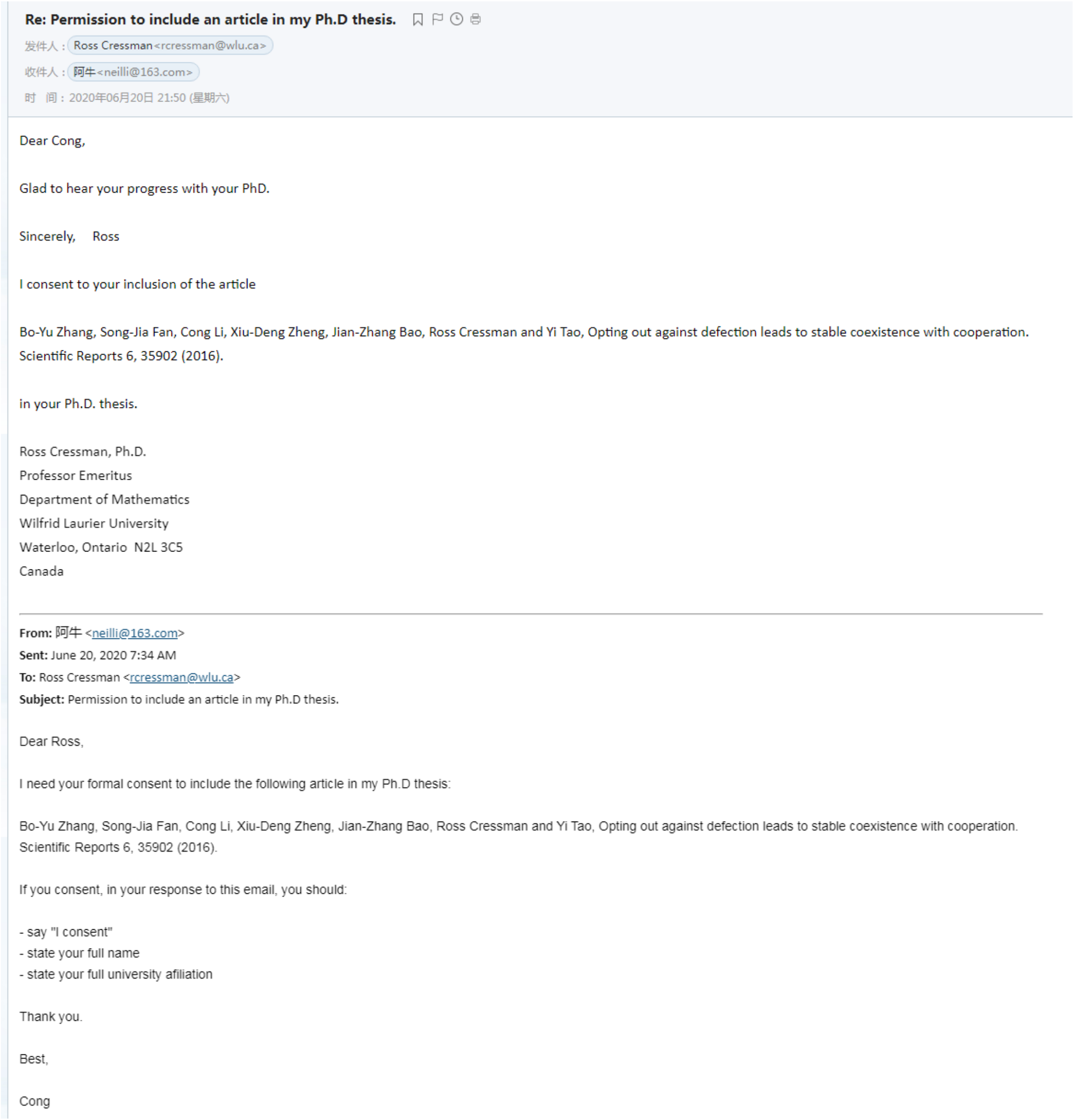
1. Identification of the coauthor (complete name and location)

   Ross Cressman (Professor Emeritus)
   Department of Mathematics
   Wilfrid Laurier University
   Waterloo, Ontario N2L 3C5
   Canada

2. Identification of the articles

   Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. *Scientific Reports* **6**, 35902 (2016).

3. Declaration



Re: Permission to include an article in my Ph.D thesis.

发件人：Ross Cressman<rcressman@wlu.ca>

收件人：阿牛<neilli@163.com>

时 间：2020年06月20日 21:50 (星期六)

Dear Cong,

Glad to hear your progress with your PhD.

Sincerely, Ross

I consent to your inclusion of the article

Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. Scientific Reports 6, 35902 (2016).

in your Ph.D. thesis.

Ross Cressman, Ph.D.
Professor Emeritus
Department of Mathematics
Wilfrid Laurier University
Waterloo, Ontario N2L 3C5
Canada

From: 阿牛 <neilli@163.com>
Sent: June 20, 2020 7:34 AM
To: Ross Cressman <rcressman@wlu.ca>
Subject: Permission to include an article in my Ph.D thesis.

Dear Ross,

I need your formal consent to include the following article in my Ph.D thesis:

Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. Scientific Reports 6, 35902 (2016).

If you consent, in your response to this email, you should:

- say "I consent"
- state your full name
- state your full university afiliation

Thank you.

Best,

Cong

PERMISSION FROM THE EDITOR OF A JOURNAL TO INCLUDE ARTICLES IN A DOCTORAL THESIS

1. Identification of the journal

   Physical Review E, Physical Review Letters

2. Identification of editor (complete name and location)

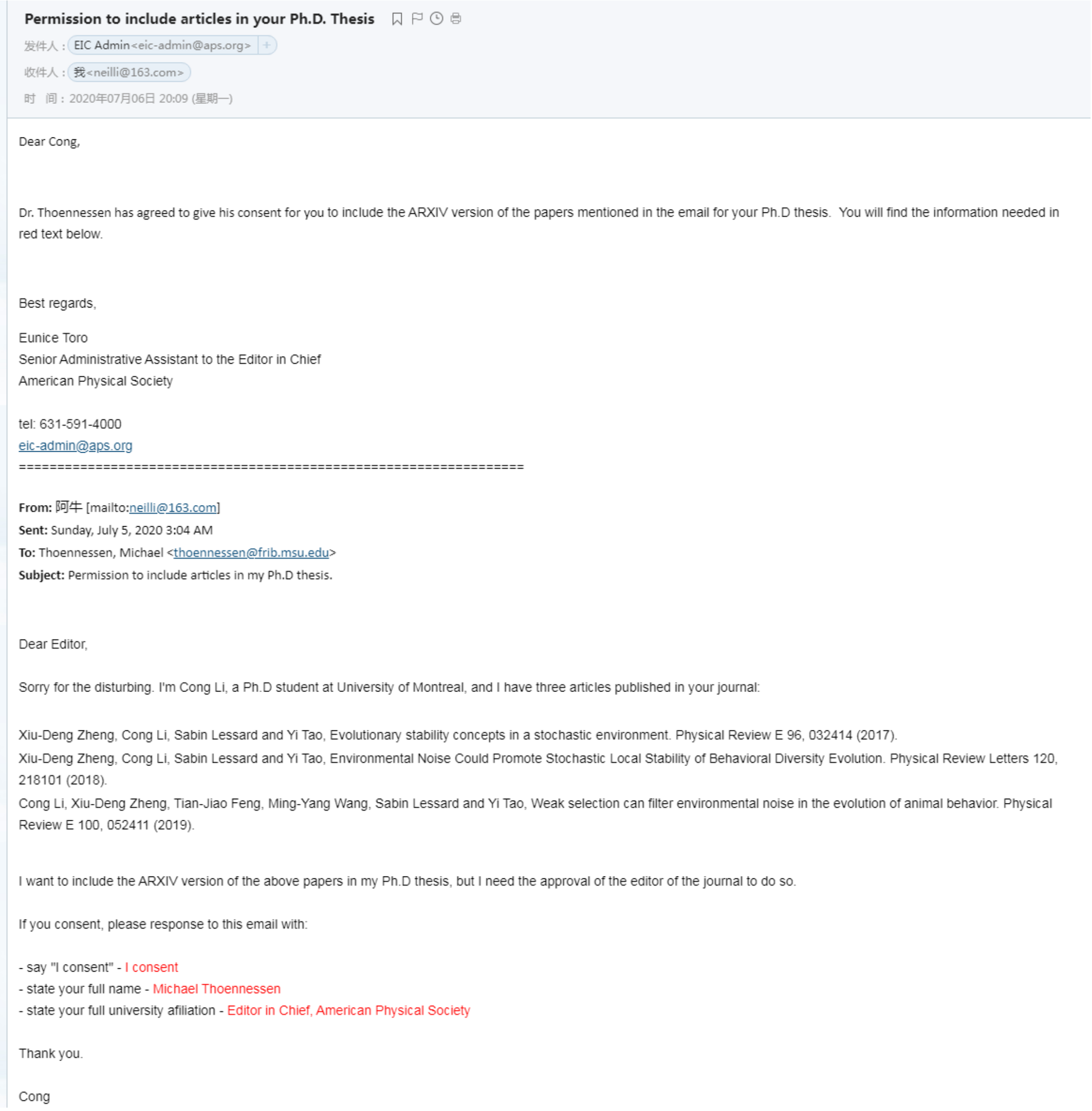
   Michael Thoennessen (Editor in Chief, American Physical Society)
   1 Research Road Ridge, NY 11961-2701, USA

3. Identification of the articles

   Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Evolutionary stability concepts in a stochastic environment. *Physical Review E* **96**, 032414 (2017).

   Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Environmental Noise Could Promote Stochastic Local Stability of Behavioral Diversity Evolution. *Physical Review Letters* **120**, 218101 (2018).

   Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. *Physical Review E* **100**, 052411 (2019).



Permission to include articles in your Ph.D. Thesis

发件人：EIC Admin<eic-admin@aps.org>

收件人：我<neilli@163.com>

时　间：2020年07月06日 20:09 (星期一)

Dear Cong,

Dr. Thoennessen has agreed to give his consent for you to include the ARXIV version of the papers mentioned in the email for your Ph.D thesis. You will find the information needed in red text below.

Best regards,

Eunice Toro
Senior Administrative Assistant to the Editor in Chief
American Physical Society

tel: 631-591-4000
eic-admin@aps.org
===

From: 阿牛 [mailto:neilli@163.com]
Sent: Sunday, July 5, 2020 3:04 AM
To: Thoennessen, Michael <thoennessen@frib.msu.edu>
Subject: Permission to include articles in my Ph.D thesis.

Dear Editor,

Sorry for the disturbing. I'm Cong Li, a Ph.D student at University of Montreal, and I have three articles published in your journal:

Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Evolutionary stability concepts in a stochastic environment. Physical Review E 96, 032414 (2017).
Xiu-Deng Zheng, Cong Li, Sabin Lessard and Yi Tao, Environmental Noise Could Promote Stochastic Local Stability of Behavioral Diversity Evolution. Physical Review Letters 120, 218101 (2018).
Cong Li, Xiu-Deng Zheng, Tian-Jiao Feng, Ming-Yang Wang, Sabin Lessard and Yi Tao, Weak selection can filter environmental noise in the evolution of animal behavior. Physical Review E 100, 052411 (2019).

I want to include the ARXIV version of the above papers in my Ph.D thesis, but I need the approval of the editor of the journal to do so.

If you consent, please response to this email with:

- say "I consent" - I consent
- state your full name - Michael Thoennessen
- state your full university afiliation - Editor in Chief, American Physical Society

Thank you.

Cong

PERMISSION FROM THE PUBLISHER OF JOURNALS TO INCLUDE ARTICLES IN A DOCTORAL THESIS

1. Identification of the journals

   Theoretical Population Biology
   Journal of Theoretical Biology

2. Identification of publisher (name and location)

   Kaveri Thakuria (Permissions Helpdesk of Elsevier)
   The Boulevard, Langford Lane,
   Kidlington, Oxford, OX5 1GB,
   United Kingdom

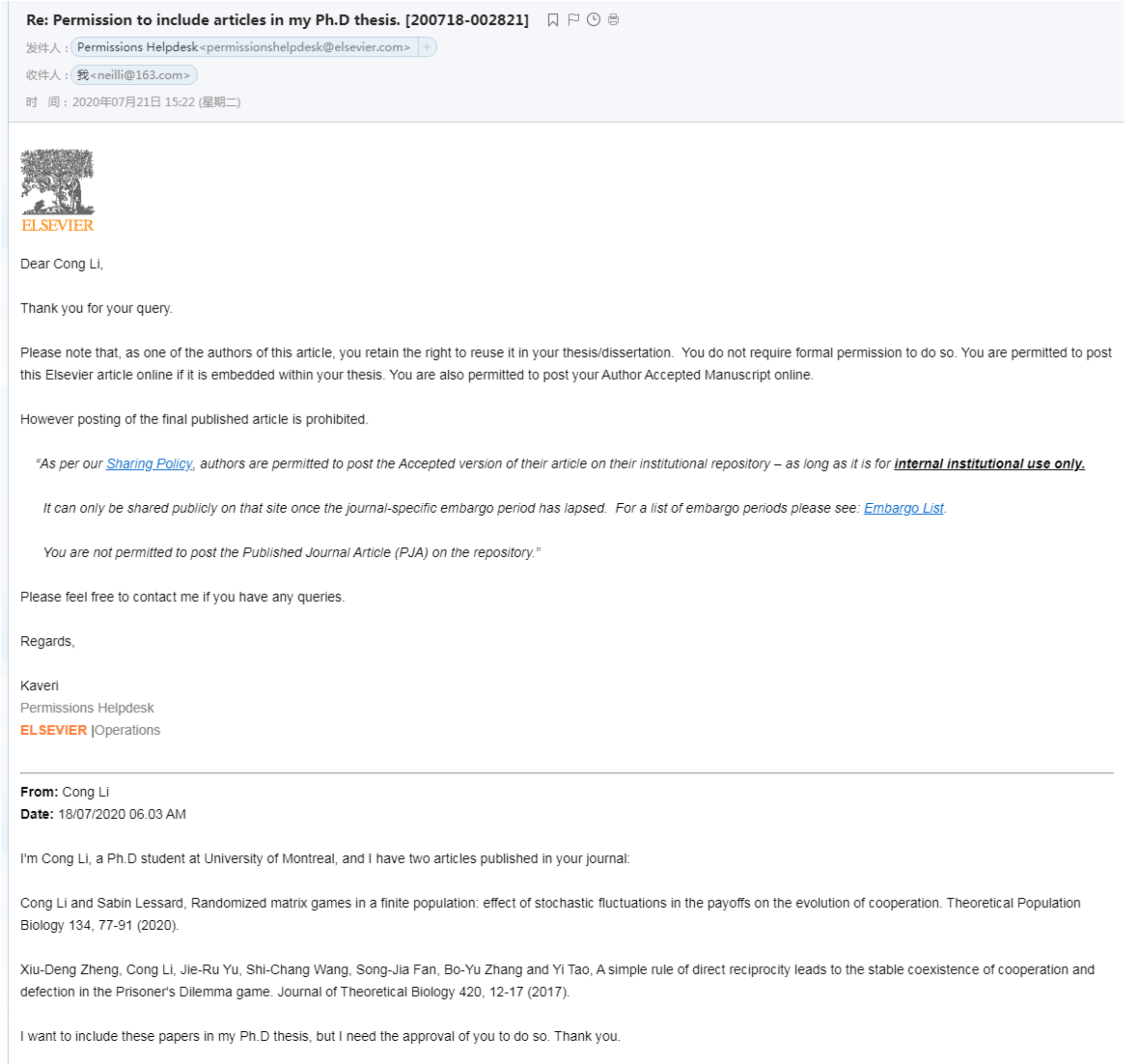
3. Identification of the articles

   Cong Li, Ting Ji, Qiao-qiao He, Xiu-deng Zheng, Sabin Lessard and Yi Tao, Uncertainty in payoffs for defection could be conducive to the evolution of cooperation. Under review

   Cong Li and Sabin Lessard, Randomized matrix games in a finite population: effect of stochastic fluctuations in the payoffs on the evolution of cooperation. accept by *Theoretical Population Biology* **134**, 77-91 (2020).

   Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. *Journal of Theoretical Biology* **420**, 12-17 (2017).

   Cong Li and Sabin Lessard, Effect of the opting-out strategy on conditions for selection to favor the evolution of cooperation in a finite population. Under review.



Re: Permission to include articles in my Ph.D thesis. [200718-002821]

发件人：Permissions Helpdesk<permissionshelpdesk@elsevier.com>

收件人：我<neilli@163.com>

时 间：2020年07月21日 15:22 (星期二)

ELSEVIER

Dear Cong Li,

Thank you for your query.

Please note that, as one of the authors of this article, you retain the right to reuse it in your thesis/dissertation. You do not require formal permission to do so. You are permitted to post this Elsevier article online if it is embedded within your thesis. You are also permitted to post your Author Accepted Manuscript online.

However posting of the final published article is prohibited.

> *"As per our Sharing Policy, authors are permitted to post the Accepted version of their article on their institutional repository – as long as it is for **internal institutional use only.***
>
> *It can only be shared publicly on that site once the journal-specific embargo period has lapsed. For a list of embargo periods please see: Embargo List.*
>
> *You are not permitted to post the Published Journal Article (PJA) on the repository."*

Please feel free to contact me if you have any queries.

Regards,

Kaveri
Permissions Helpdesk
ELSEVIER |Operations

From: Cong Li
Date: 18/07/2020 06.03 AM

I'm Cong Li, a Ph.D student at University of Montreal, and I have two articles published in your journal:

Cong Li and Sabin Lessard, Randomized matrix games in a finite population: effect of stochastic fluctuations in the payoffs on the evolution of cooperation. Theoretical Population Biology 134, 77-91 (2020).

Xiu-Deng Zheng, Cong Li, Jie-Ru Yu, Shi-Chang Wang, Song-Jia Fan, Bo-Yu Zhang and Yi Tao, A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. Journal of Theoretical Biology 420, 12-17 (2017).

I want to include these papers in my Ph.D thesis, but I need the approval of you to do so. Thank you.

PERMISSION FROM THE EDITOR OF A JOURNAL TO INCLUDE ARTICLES IN A DOCTORAL THESIS

1. Identification of the journal

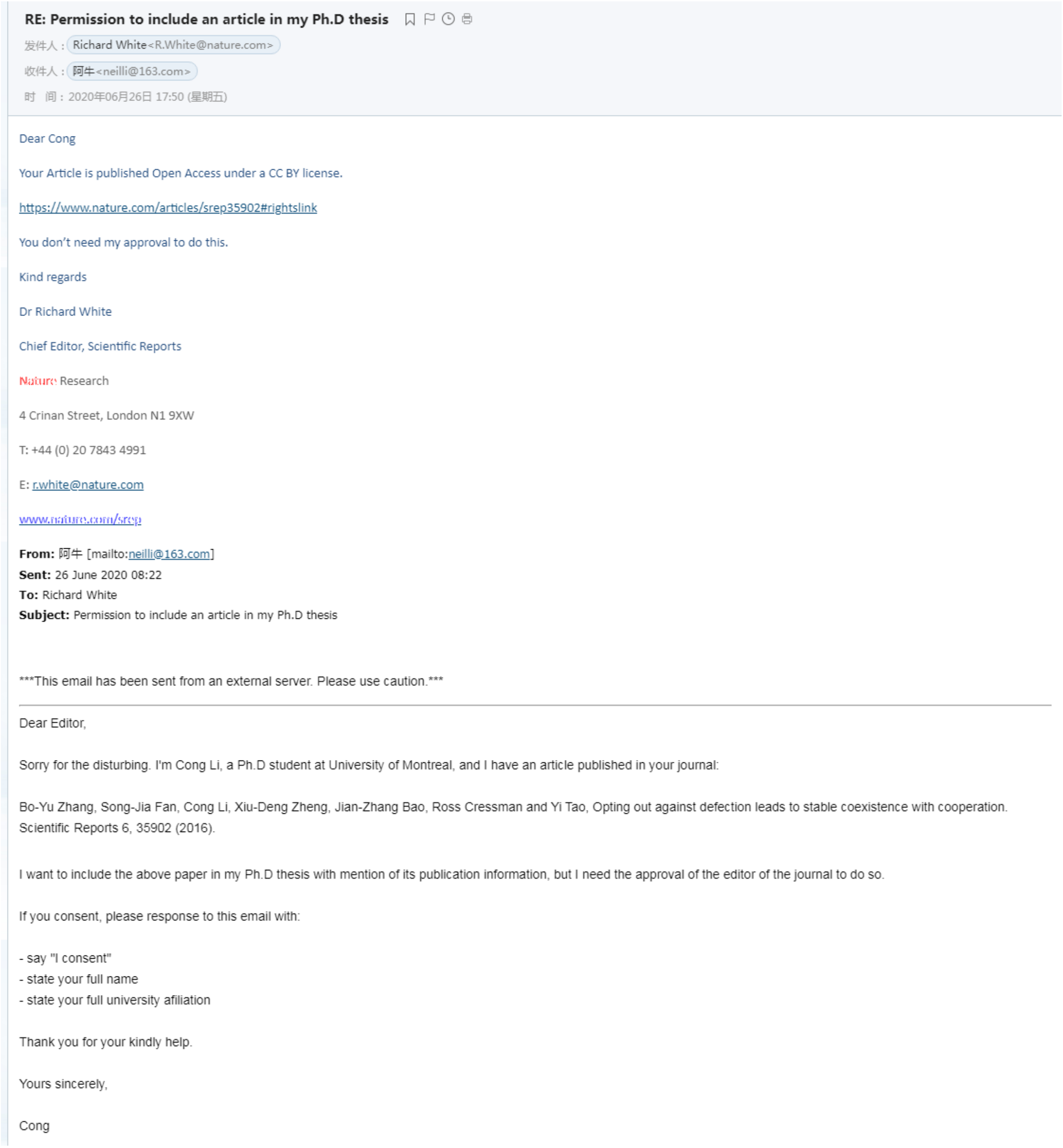
   Scientific Reports

2. Identification of editor (complete name and location)

   Richard White (Chief Editor, Scientific Reports)
   4 Crinan Street, London N1 9XW
   United Kingdom

3. Identification of the articles

   Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. *Scientific Reports* **6**, 35902 (2016).



RE: Permission to include an article in my Ph.D thesis

发件人：Richard White<R.White@nature.com>

收件人：阿牛<neilli@163.com>

时　间：2020年06月26日 17:50 (星期五)

Dear Cong

Your Article is published Open Access under a CC BY license.

https://www.nature.com/articles/srep35902#rightslink

You don't need my approval to do this.

Kind regards

Dr Richard White

Chief Editor, Scientific Reports

Nature Research

4 Crinan Street, London N1 9XW

T: +44 (0) 20 7843 4991

E: r.white@nature.com

www.nature.com/srep

From: 阿牛 [mailto:neilli@163.com]
Sent: 26 June 2020 08:22
To: Richard White
Subject: Permission to include an article in my Ph.D thesis

This email has been sent from an external server. Please use caution.

Dear Editor,

Sorry for the disturbing. I'm Cong Li, a Ph.D student at University of Montreal, and I have an article published in your journal:

Bo-Yu Zhang, Song-Jia Fan, Cong Li, Xiu-Deng Zheng, Jian-Zhang Bao, Ross Cressman and Yi Tao, Opting out against defection leads to stable coexistence with cooperation. Scientific Reports 6, 35902 (2016).

I want to include the above paper in my Ph.D thesis with mention of its publication information, but I need the approval of the editor of the journal to do so.

If you consent, please response to this email with:

- say "I consent"
- state your full name
- state your full university afiliation

Thank you for your kindly help.

Yours sincerely,

Cong