Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2007, Article ID 82184, 34 pages
doi:10.1155/2007/82184

Research Article

Rational Probabilistic Deciders—Part II: Collective Behavior

P. T. Kabamba, W.-C. Lin, and S. M. Meerkov

Received 7 December 2006; Accepted 23 February 2007

Recommended by Jingshan Li

This paper explores the behavior of rational probabilistic deciders (RPDs) in three types of collectives: zero sum matrix games, fractional interactions, and Edgeworth exchange economies. The properties of steady states and transients are analyzed as a function of the level of rationality, N, and, in some cases, the size of the collective, M. It is shown that collectives of RPDs, may or may not behave rationally, depending, for instance, on the relationship between N and M (under fractional interactions) or N and the minimum amount of product exchange (in Edgeworth economies). The results obtained can be useful for designing rational reconfigurable systems that can autonomously adapt to changing environments.

1. Introduction

1.1. Issues addressed and results obtained. The notion of a rational probabilistic decider (RPD) was introduced in [1]. Roughly speaking, an RPD is a stochastic system, which takes less penalized decisions with larger probabilities than other ones (see Section 1.2 below for a precise definition). Two types of RPDs have been analyzed: local (L-RPD) and global (G-RPD). L-RPDs take their decisions based on the penalty function of their current states, while G-RPDs consider penalties of other states as well.

In [1], the behavior of individual RPDs was investigated. It was shown that asymptotic properties of both L- and G-RPDs are the same: both converge to the best decision when the so-called level of rationality tends to infinity. However, their temporal properties are different: G-RPDs react at a faster rate, which gives them an advantage in nonstationary environments.

The current paper is devoted to group (or collective) behavior of RPDs. A collective of RPDs is formed by assuming that their penalty functions are interrelated in the sense that the penalty of each depends on the actions of the others. Three types of penalty functions are considered. In the first one, the penalty function is defined by the payoff matrix of a zero sum game (with the players being RPDs). In the second, referred to as fractional, the penalty function depends on the fraction of the group members that select a particular decision. The fractional interactions considered are of two types: homogeneous and nonhomogeneous. In the homogeneous case, all group members are penalized identically, while in the nonhomogeneous one the penalty depends on the particular subgroup to which an RPD belongs. Finally, the third type of penalty function is defined by an economic model referred to as the Edgeworth exchange economy.

In all these types of interactions, the question of interest is that will a collective of RPDs behave rationally, that is, converge to the state where the penalty is minimized? Analyzing this question, this paper reports the following results.

In the matrix game environment,

(a) both L- and G-RPDs converge to the min-max point if the payoff matrix has a saddle in pure strategies; this result is analogous to that obtained in [2], where rational behavior was modeled by finite automata;

(b) if the saddle point is in mixed strategies, both L- and G-RPDs are unable to find these, however, G-RPDs playing against L-RPDs win by converging to the upper value of the game; this result is novel;

(c) if an L-RPD or G-RPD is playing against a human that uses his mixed optimal strategy, the RPD is able to find its mixed optimal strategy provided that the payoff matrix is symmetric; this is different from [3] in that finite automata cannot find mixed optimal strategies when playing against humans;

(d) rates of convergence for G-RPDs are faster than those for L-RPDs, giving G-RPDs an advantage in the transients when playing against L-RPDs in games with a saddle in pure strategies; this result is novel—the previous literature did not address this issue.

Under homogeneous fractional interaction,

(a) a collective behaves optimally if the level of rationality of each RPD grows at least as fast as the size of the collective; this is similar to the result obtained in [4, 5], where rational behavior was modeled by finite automata and general dynamical systems, respectively;

(b) although G-RPDs behave similarly to L-RPDs in the steady state, the rate of convergence for G-RPDs is much faster than that of L-RPDs; this result is also novel.

Under nonhomogeneous fractional interaction,

(a) a collective behaves optimally even if the size of the collective tends to infinity as long as the level of rationality of each individual is sufficiently large; this result is similar to that obtained in [5];

(b) as in homogeneous fractional interactions, the rate of convergence for G-RPDs is much faster than that of L-RPDs, which is also a new result.

In the Edgeworth exchange economy [6],

(a) G-RPDs with an identical level of rationality converge to a particular Pareto equilibrium, irrespective of the initial product allocation; this result is different from the classical one where the convergence is to a subset of the Pareto equilibrium, which is defined by the initial allocation; this result is novel;
(b) if the level of rationality of the two G-RPDs are not identical, the resulting stable Pareto equilibrium gives advantage to the one with larger rationality; this result is also novel.

1.2. Definition of RPD. To make this paper self-contained, below we briefly recapitulate the definition of RPDs; for more details, see [1], where a comprehensive literature review is also included.

A *probabilistic decider* (PD) is a stochastic system defined by a quadruple,

$$(\mathcal{S},\Phi,N,\mathcal{P}), \tag{1.1}$$

where $\mathcal{S}=\{1,2,\ldots,s\}$ is the decision space; $\Phi=[\varphi_1,\varphi_2,\ldots,\varphi_s]$ is the penalty function; $N\in(0,\infty)$ is the level of rationality; and $\mathcal{P}=\{P_1,P_2,\ldots,P_s\}$ is a set of transition probabilities such that the probability of a state transition from state $i\in\mathcal{S}$ is

$$P[x(n+1)\neq i \mid x(n)=i]=P_i(\varphi_1,\varphi_2,\ldots,\varphi_s;N)=P_i(\Phi;N),\quad n=0,1,2,\ldots. \tag{1.2}$$

When a state transition occurs, all other states are selected equiprobably, that is,

$$P[x(n+1)=j \mid x(n)=i]=\frac{P_i(\Phi;N)}{s-1}\quad \text{for } j\neq i. \tag{1.3}$$

Let $\kappa_i(\Phi;N)$ denote the steady state probability of state $i\in\mathcal{S}$ when Φ is constant (i.e., the environment is stationary). A PD is rational (i.e., RPD) if the following takes place: inequality $\varphi_i<\varphi_j$ implies that

$$\frac{\kappa_i}{\kappa_j}>1,\quad \forall N\in(0,\infty), \tag{1.4}$$

and, moreover,

$$\frac{\kappa_i}{\kappa_j}\longrightarrow\infty\quad \text{as } N\longrightarrow\infty. \tag{1.5}$$

An RPD is local (i.e., L-RPD) if

$$P_i(\Phi;N)=P^{I}(\varphi_i,N),\quad i\in\mathcal{S}, \tag{1.6}$$

that is, L-RPDs take decisions based on the penalty of the current state. An RPD is global (i.e., G-RPD) if $\mathcal{S}=\{1,2\}$,

$$\frac{P_1(\Phi;N)}{P_2(\Phi;N)}=P^{II}(\varphi_1,\varphi_2;N), \tag{1.7}$$

that is, G-RPDs take decisions based on the penalties of all states. The properties of P^I and P^{II} are described in [1]. In particular, P^{II} may be of the form

$$P^{II} = \frac{F(G(N,\varphi_1/\varphi_2))}{F(G(N,\varphi_2/\varphi_1))}, \tag{1.8}$$

where functions F and G are characterized in [1]. Examples of appropriate functions F and G can be given as follows:

$$F(x) = \frac{x}{1+x}, \qquad G(N,y) = y^N. \tag{1.9}$$

The current paper addresses the issue of *collective behavior* of M RPDs, that is, when the penalty function φ_i, $i = 1,2,\ldots,s$, is not constant but is changing in accordance with changing states of all members of the collective.

1.3. Paper outline. The outline of this paper is as follows. In Section 2, we introduce the notion of a collective of RPDs and describe the problems addressed. The collective behaviors of RPDs in zero sum matrix games, under fractional interactions, and in Edgeworth exchange economies are investigated in Sections 3–5, respectively, and Section 6 gives conclusions. All proofs are given in the appendices.

2. Collective of RPDs

2.1. Modeling. A *collective* of RPDs is defined as a set of RPDs, where the penalties incurred by an RPD depend not only on the its state but also on the states of the other RPDs. Specifically, consider a set of M RPDs. Denote the jth RPD by the quadruple, $(\mathcal{S}^j,\Phi^j,N^j,\mathcal{P}^j)$, $j = 1,2,\ldots,M$, where

(a) $\mathcal{S}^j = \{x_1^j,x_2^j,\ldots,x_{s_j}^j\}$ is the decision space of the jth RPD. At each time moment, $n = 0,1,2,\ldots$, the RPD is in one of the states of $\mathcal{S}^j$;

(b) $\Phi^j(n) = [\varphi_1^j(n) \quad \varphi_2^j(n) \quad \cdots \quad \varphi_{s_j}^j(n)]$ is a vector, where $\varphi_i^j(n)$ denotes the penalty associated with state $x_i^j \in \mathcal{S}^j$ at time n. Furthermore, we assume

$$\varphi_i^j(n) = \phi^j(x^1(n),x^2(n),\ldots,x^{j-1}(n),x_i^j,x^{j+1}(n),\ldots,x^M(n)), \tag{2.1}$$

where $x^k(n) \in \mathcal{S}^k$, $k \neq j$, denotes the state of the kth RPD at time n;

(c) $N^j \in (0,\infty)$ is a positive number, which denotes the level of rationality of the jth RPD;

(d) $\mathcal{P}^j = \{P_1^j(\Phi^j(n),N^j),P_2^j(\Phi^j(n),N^j),\ldots,P_{s_j}^j(\Phi^j(n),N^j)\}$, where

$$0 < P_i^j(\Phi^j(n),N^j) < 1, \tag{2.2}$$

is a set of transition probabilities depending on $\Phi^j(n)$ and N^j.

The collective of RPDs can operate in the following two modes.

(α) *Parallel operation:* if at time n the jth RPD, $j = 1,2,\ldots,M$, is in state x_i^j, then the probability that it will make a state transition at time $n+1$ is

$$P[x^j(n+1) \neq x_i^j \mid x^j(n) = x_i^j] = P_i^j(\Phi^j(n);N^j). \tag{2.3}$$

When a state transition occurs, the RPD chooses any other state with equal probability, that is,

$$P[x^j(n+1) = x_l^j \mid x^j(n) = x_i^j] = \frac{P_i^j(\Phi^j(n);N^j)}{s^j - 1} \quad \text{for } x_i^j \neq x_l^j. \tag{2.4}$$

(β) *Sequential operation:* at each time n, one of the RPDs is chosen with probability $1/M$. Suppose the jth RPD is chosen and that it is in state x_i^j. Then, at time $n+1$, it will make a state transition according to (2.3) and (2.4), while all other RPDs remain in their original states.

2.2. Problems. The interactions among the RPDs in a collective are described by the penalties in (2.1), and are defined by the environment surrounding the collective. In this paper, the behavior of collectives of RPDs in zero sum matrix games, under fractional interactions, and in Edgeworth exchange economies are considered. In particular, we address the following problems: given a collective of RPDs and an environment,

(i) analyze the steady state probabilities of various decisions as a function of the level of rationality and the parameters of the environment;

(ii) investigate the rates of convergence to the steady state.

Exact formulations and solutions of these problems are given in Sections 3–5.

3. Collective of RPDs in zero sum matrix games

3.1. Environment and steady state probabilities. Consider a 2×2 zero sum matrix game with payoff matrix

$$M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}, \tag{3.1}$$

where m_{kl}, $k,l = 1,2$, is the payoff to the first player when it selects action k and the second selects action l. Without loss of generality, we assume that

$$-0.95 \leq m_{kl} \leq 0.95 \quad k,l = 1,2. \tag{3.2}$$

The two players of the matrix game form a collective of RPDs described below.

(i) The first and second player of the matrix game are the first and second RPD of the collective, respectively.

(ii) Let $\mathcal{S}^j = \{1,2\}$, $j = 1,2$, where the states in $\mathcal{S}^j$ correspond to the actions of the RPDs.

(iii) Converting payoffs to penalties, let

$$\Phi^j(n) = \left[\varphi_1^j(n) \quad \varphi_2^j(n)\right], \tag{3.3}$$

where

$$\varphi_i^j(n) = \begin{cases} \dfrac{1-m_{il}}{2}, & \text{if } j=1 \text{ and the second RPD selects state } l \text{ at time } n, \\ \dfrac{1+m_{ki}}{2}, & \text{if } j=2 \text{ and the first RPD selects state } k \text{ at time } n. \end{cases} \tag{3.4}$$

(iv) The RPDs of the collective are L-RPDs or G-RPDs. If the jth RPD is L-RPD, then

$$P_i^j(\Phi^j(n);N^j) = \left(\varphi_i^j(n)\right)^{N^j} \quad \text{for } i=1,2. \tag{3.5}$$

If the jth RPD is G-RPD with functions F and G given by (1.9), then

$$P_i^j(\Phi^j(n);N^j) = \frac{\left(\varphi_i^j(n)\right)^{N^j}}{\left(\varphi_1^j(n)\right)^{N^j} + \left(\varphi_2^j(n)\right)^{N^j}} \quad \text{for } i=1,2. \tag{3.6}$$

(v) The RPDs make state transitions according to parallel operation mode.

Due to assumptions (i)–(v), the dynamics of a collective of RPDs playing the 2×2 matrix game is described by an ergodic Markov chain with transition matrix

$$A = \begin{bmatrix} (1-a_{11})(1-b_{11}) & (1-a_{11})b_{11} & a_{11}(1-b_{11}) & a_{11}b_{11} \\ (1-a_{12})b_{12} & (1-a_{12})(1-b_{12}) & a_{12}b_{12} & a_{12}(1-b_{12}) \\ a_{21}(1-b_{21}) & a_{21}b_{21} & (1-a_{21})(1-b_{21}) & (1-a_{21})b_{21} \\ a_{22}b_{22} & a_{22}(1-b_{22}) & (1-a_{22})b_{22} & (1-a_{22})(1-b_{22}) \end{bmatrix}, \tag{3.7}$$

where

$$a_{kl} = \begin{cases} \left(\dfrac{1-m_{kl}}{2}\right)^{N^1}, & \text{if the first RPD is an L-RPD}, \\ \dfrac{((1-m_{kl})/2)^{N^1}}{((1-m_{1l})/2)^{N^1} + ((1-m_{2l})/2)^{N^1}}, & \text{if the first RPD is a G-RPD}, \end{cases} \tag{3.8}$$

$$b_{kl} = \begin{cases} \left(\dfrac{1+m_{kl}}{2}\right)^{N^2}, & \text{if the second RPD is an L-RPD}, \\ \dfrac{((1+m_{kl})/2)^{N^2}}{((1+m_{k1})/2)^{N^2} + ((1+m_{k2})2)^{N^2}}, & \text{if the second RPD is a G-RPD}. \end{cases} \tag{3.9}$$

Let $\kappa = [\kappa_{11} \quad \kappa_{12} \quad \kappa_{21} \quad \kappa_{22}]$ be a row vector, where κ_{kl} is the steady state probability of the first RPD selecting state k and the second selecting l. Then, κ can be calculated from the equations

$$\kappa = \kappa A, \quad \sum_{kl} \kappa_{kl} = 1. \tag{3.10}$$

The solutions to (3.10) are given by

$$\kappa_{11} = -\frac{\Delta_{11}}{\Delta}, \qquad \kappa_{12} = -\frac{\Delta_{12}}{\Delta}, \qquad \kappa_{21} = -\frac{\Delta_{21}}{\Delta}, \qquad \kappa_{22} = -\frac{\Delta_{22}}{\Delta}, \tag{3.11}$$

where

$$\begin{aligned}
\Delta = {} & -a_{12}a_{21}b_{11} - a_{21}a_{22}b_{11} - a_{11}a_{22}b_{12} - a_{21}a_{22}b_{12} - a_{11}a_{21}b_{11}b_{12} \\
& + a_{12}a_{21}b_{11}b_{12} + a_{11}a_{22}b_{11}b_{12} - a_{12}a_{22}b_{11}b_{12} - a_{11}a_{12}b_{21} - a_{11}a_{22}b_{21} \\
& - a_{12}b_{11}b_{21} + a_{11}a_{12}b_{11}b_{21} + a_{12}a_{21}b_{11}b_{21} - a_{22}b_{11}b_{21} + a_{11}a_{22}b_{11}b_{21} \\
& + a_{21}a_{22}b_{11}b_{21} - a_{11}b_{12}b_{21} + a_{11}a_{12}b_{12}b_{21} + a_{11}a_{21}b_{12}b_{21} - a_{22}b_{12}b_{21} \\
& + a_{12}a_{22}b_{12}b_{21} + a_{21}a_{22}b_{12}b_{21} - a_{11}a_{12}b_{22} - a_{12}a_{21}b_{22} - a_{12}b_{11}b_{22} \\
& + a_{11}a_{12}b_{11}b_{22} - a_{21}b_{11}b_{22} + a_{11}a_{21}b_{11}b_{22} + a_{12}a_{22}b_{11}b_{22} \\
& + a_{21}a_{22}b_{11}b_{22} - a_{11}b_{12}b_{22} + a_{11}a_{12}b_{12}b_{22} - a_{21}b_{12}b_{22} \\
& + a_{12}a_{21}b_{12}b_{22} + a_{11}a_{22}b_{12}b_{22} + a_{21}a_{22}b_{12}b_{22} - a_{11}a_{21}b_{21}b_{22} \\
& + a_{12}a_{21}b_{21}b_{22} + a_{11}a_{22}b_{21}b_{22} - a_{12}a_{22}b_{21}b_{22}, \\
\Delta_{11} = {} & a_{21}a_{22}b_{12} + a_{22}b_{12}b_{21} - a_{12}a_{22}b_{12}b_{21} - a_{21}a_{22}b_{12}b_{21} \\
& + a_{12}a_{21}b_{22} + a_{21}b_{12}b_{22} - a_{12}a_{21}b_{12}b_{22} - a_{21}a_{22}b_{12}b_{22} \\
& - a_{12}a_{21}b_{21}b_{22} + a_{12}a_{22}b_{21}b_{22}, \\
\Delta_{12} = {} & a_{21}a_{22}b_{11} + a_{11}a_{22}b_{21} + a_{22}b_{11}b_{21} - a_{11}a_{22}b_{11}b_{21} \\
& - a_{21}a_{22}b_{11}b_{21} + a_{21}b_{11}b_{22} - a_{11}a_{21}b_{11}b_{22} - a_{21}a_{22}b_{11}b_{22} \\
& + a_{11}a_{21}b_{21}b_{22} - a_{11}a_{22}b_{21}b_{22}, \\
\Delta_{21} = {} & a_{11}a_{22}b_{12} - a_{11}a_{22}b_{11}b_{12} + a_{12}a_{22}b_{11}b_{12} + a_{11}a_{12}b_{22} \\
& + a_{12}b_{11}b_{22} - a_{11}a_{12}b_{11}b_{22} - a_{12}a_{22}b_{11}b_{22} + a_{11}b_{12}b_{22} \\
& - a_{11}a_{12}b_{12}b_{22} - a_{11}a_{22}b_{12}b_{22}, \\
\Delta_{22} = {} & a_{12}a_{21}b_{11} + a_{11}a_{21}b_{11}b_{12} - a_{12}a_{21}b_{11}b_{12} + a_{11}a_{12}b_{21} \\
& + a_{12}b_{11}b_{21} - a_{11}a_{12}b_{11}b_{21} - a_{12}a_{21}b_{11}b_{21} + a_{11}b_{12}b_{21} \\
& - a_{11}a_{12}b_{12}b_{21} - a_{11}a_{21}b_{12}b_{21}.
\end{aligned} \tag{3.12}$$

These expressions are used below to analyze steady states of collectives where payoff matrices lead to either pure or mixed optimal strategies.

3.2. Zero sum matrix games having pure optimal strategies. In this subsection, it is assumed that the matrix game at hand has a pure optimal strategy. Without loss of generality, assume the payoff matrix in (3.1) satisfies the relation

$$m_{21} < m_{11} < m_{12}, \tag{3.13}$$

that is, m_{11} is the saddle point, and the optimal strategy is for both RPDs to select action 1. Furthermore, assume the RPDs have the same level of rationality, that is,

$$N^1 = N^2 = N, \tag{3.14}$$

where $N \in (0, \infty)$.

3.2.1. Steady state behavior. The following analysis question is addressed.

A1: can RPDs, playing the above matrix game, find their pure optimal strategies, that is, $\kappa_{11} \to 1$ as $N \to \infty$?

Specifically, the following collectives of RPDs are of interest.

(C1) Both RPDs are L-RPDs.

(C2) Both RPDs are G-RPDs.

(C3) The first RPD is a G-RPD and the second is an L-RPD.

Evaluation of (3.11) shows that for the collectives (C1)–(C3), κ_{11} approaches 1 as N approaches infinity, that is, the RPDs converge to the saddle point reliably as N becomes arbitrarily large. This is illustrated in Figures 3.1 and 3.2 for payoff matrices,

$$\begin{bmatrix} -0.1 & -0.05 \\ -0.15 & -0.9 \end{bmatrix}, \tag{3.15}$$

$$\begin{bmatrix} 0.1 & 0.5 \\ 0.05 & -0.15 \end{bmatrix}, \tag{3.16}$$

respectively. For the matrix game (3.15) and for the collectives (C1) and (C2), the values of N that are required to converge reliably to the saddle point, for example, $\kappa_{11} = 0.95$, are 76 and 87, respectively. For the matrix game (3.16), the required values of N are 20 and 68, respectively. Based on the above, the following observations can be made.

(a) Although a G-RPD uses more information than an L-RPD, this does not lead to an advantage as $N \to \infty$ (in the sense that the G-RPD does not receive more than the optimal payoff).

(b) Surprisingly, the required N for reliable selection of the saddle point is larger when both players are G-RPDs than when both are L-RPDs. In some games, as the one with payoff matrix (3.16), the difference is quite large.

3.2.2. Transient behavior. From the above analysis, G-RPDs do not outperform L-RPDs in the steady state. However, the fact that G-RPDs use more information should, in some way, give G-RPDs advantage over L-RPDs. Hence, the following question is addressed.

A2: can G-RPDs outperform L-RPDs during the transients of a matrix game?

The rates of convergence in time are analyzed first. Given the payoff matrix defined in (3.15), Figure 3.3 shows the behavior of the second largest eigenvalue, λ_2, of matrix A as a

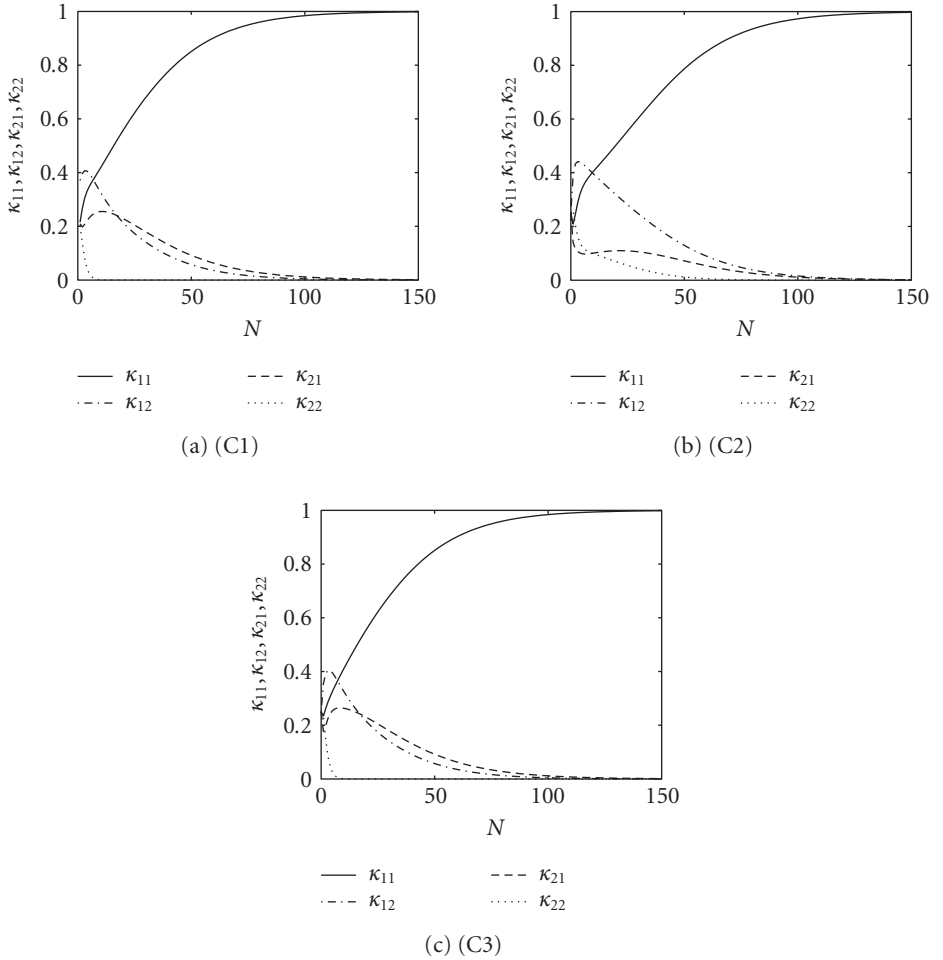

Figure 3.1. Steady state probabilities κ_{kl} versus N for the collectives (C1)–(C3) with payoff matrix (3.15).

function of N for the collectives (C1) and (C2). As it follows from this figure, it will take an arbitrarily long time for the game to converge as N becomes large if both players are L-RPDs, while this is not true if both players are G-RPDs. Moreover, when both players are G-RPDs, the time required for convergence becomes shorter as N becomes larger and, when N becomes arbitrarily large, that time tends to zero.

Next, consider a matrix game played by (C3). Let $P_G(n)$ and $P_L(n)$ denote the payoffs to the G-RPD and L-RPD, respectively, at time n, and let

$$P_G^{\text{avg}}(n) = \frac{1}{n}\sum_{i=0}^{n} P_G(i), \qquad P_L^{\text{avg}}(n) = \frac{1}{n}\sum_{i=0}^{n} P_L(i). \tag{3.17}$$

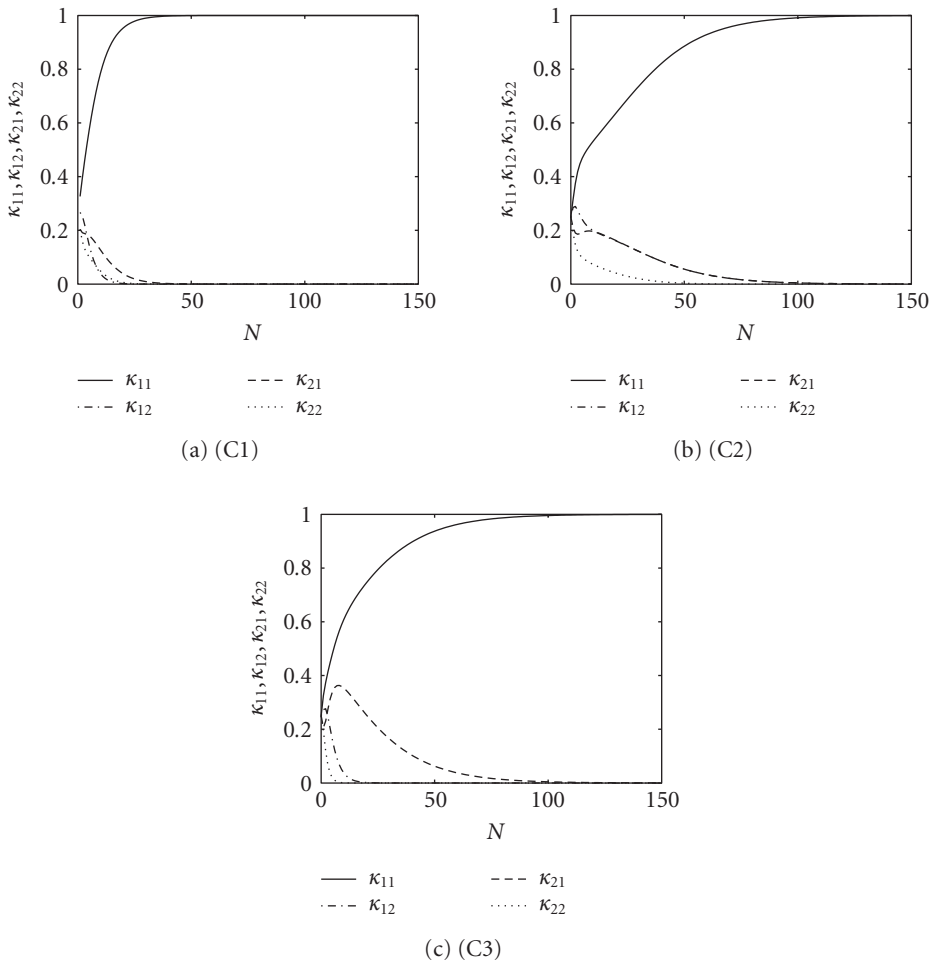

Figure 3.2. Steady state probabilities κ_{kl} versus N for the collectives (C1)–(C3) with payoff matrix (3.16).

Figure 3.4 shows $P_G^{\text{avg}}(n)$ and $P_L^{\text{avg}}(n)$ as a function of time n for the matrix game,

$$\begin{bmatrix} 0 & 0.5 \\ -0.5 & 0 \end{bmatrix}, \tag{3.18}$$

assuming $N = 5$ and players initially at the saddle point. Clearly, G-RPDs, being able to converge faster, have advantage over L-RPDs during the transients of the game.

3.3. Zero sum matrix games having mixed optimal strategies. In this subsection, it is assumed that the matrix game at hand has a mixed optimal strategy. Without loss of

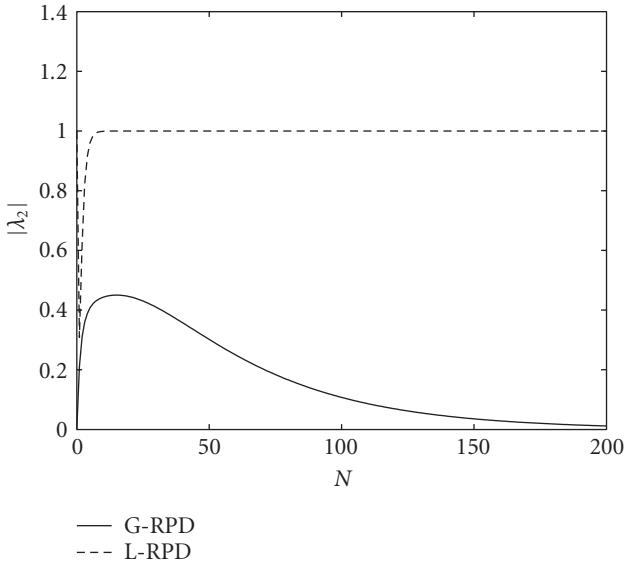

Figure 3.3. Second largest eigenvalue $|\lambda_2|$ versus N for (C1), (C2), and payoff matrix (3.15).

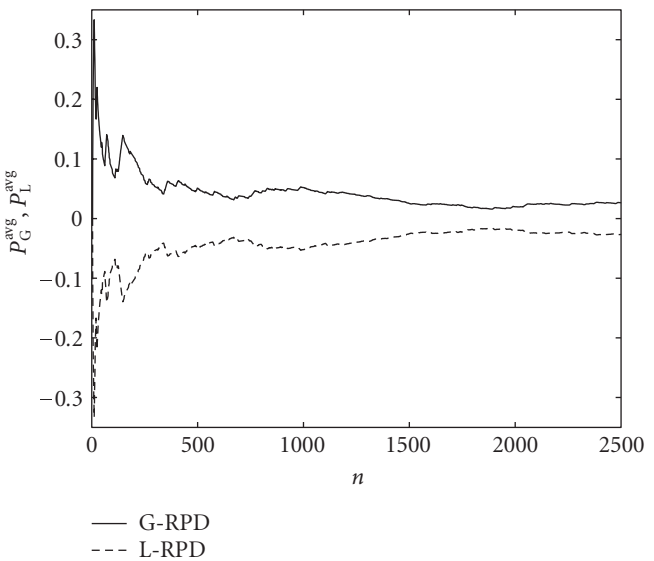

Figure 3.4. Average payoffs $P_G^{\text{avg}}(n)$ and $P_L^{\text{avg}}(n)$ as a function of n.

generality, assume the payoff matrix (3.1) satisfies the relation,

$$m_{11} \geq m_{22} > m_{12} \geq m_{21}. \tag{3.19}$$

Hence, the mixed optimal strategy is as follows: the first player selects action 1 with probability

$$\kappa^{1*} = \frac{m_{22} - m_{21}}{(m_{11} + m_{22}) - (m_{12} + m_{21})}, \tag{3.20}$$

and the second player selects action 1 with probability

$$\kappa^{2*} = \frac{m_{22} - m_{12}}{(m_{11} + m_{22}) - (m_{12} + m_{21})}. \tag{3.21}$$

The following analysis question is addressed.

A: can RPDs playing the matrix game find their mixed optimal strategies?

To answer this question, collectives (C1)–(C3) of Section 3.2.1 with $N_1 = N_2 = N$ are considered. Evaluating (3.11), one can see that none of the RPDs is able to find the mixed optimal strategy. Specifically,

(i) for (C1), as $N \to \infty$, the game value converges to either the lower or upper value of the game, depending on the payoff matrix, that is, to either m_{12} or m_{22};

(ii) for (C2), as $N \to \infty$, the game value converges to the average of the entries of the payoff matrix, that is, to $(m_{11} + m_{12} + m_{21} + m_{22})/4$;

(iii) for (C3), the outcome of the matrix game always converges to the upper value, m_{22}, of the game as $N \to \infty$; this means that, when N is sufficiently large, the G-RPD is always receiving more than the optimal payoff, and hence, has an advantage when playing against the L-RPD.

Since the players are not able to find their mixed optimal strategies when both are RPDs, we consider the following additional collectives.

(C4) The first player is an L-RPD and the second is a human playing according to his mixed optimal strategy.

(C5) The first player is a G-RPD and the second is a human playing according to his mixed optimal strategy.

For (C4) and (C5), the transition matrix A in (3.7) becomes

$$A = \begin{bmatrix} (1-a_{11})\kappa^{2*} & (1-a_{11})(1-\kappa^{2*}) & a_{11}\kappa^{2*} & a_{11}(1-\kappa^{2*}) \\ (1-a_{12})\kappa^{2*} & (1-a_{12})(1-\kappa^{2*}) & a_{12}\kappa^{2*} & a_{12}(1-\kappa^{2*}) \\ a_{21}\kappa^{2*} & a_{21}(1-\kappa^{2*}) & (1-a_{21})\kappa^{2*} & (1-a_{21})(1-\kappa^{2*}) \\ a_{22}\kappa^{2*} & a_{22}(1-\kappa^{2*}) & (1-a_{22})\kappa^{2*} & (1-a_{22})(1-\kappa^{2*}) \end{bmatrix}. \tag{3.22}$$

The steady state probabilities in (3.11) become

$$\kappa_{11} = -\frac{\Delta'_{11}}{\Delta'}, \qquad \kappa_{12} = -\frac{\Delta'_{12}}{\Delta'}, \qquad \kappa_{21} = -\frac{\Delta'_{21}}{\Delta'}, \qquad \kappa_{22} = -\frac{\Delta'_{22}}{\Delta'}, \tag{3.23}$$

where

$$\begin{aligned}
\Delta' &= -a_{12} - a_{22} - a_{11}\kappa^{2*} + a_{12}\kappa^{2*} - a_{21}\kappa^{2*} + a_{22}\kappa^{2*},\\
\Delta'_{11} &= a_{22}\kappa^{2*} + a_{21}\left(\kappa^{2*}\right)^2 - a_{22}\left(\kappa^{2*}\right)^2,\\
\Delta'_{21} &= a_{22} + a_{21}\kappa^{2*} - 2a_{22}\kappa^{2*} - a_{21}\left(\kappa^{2*}\right)^2 + a_{22}\left(\kappa^{2*}\right)^2,\\
\Delta'_{21} &= a_{12}\kappa^{2*} + a_{11}\left(\kappa^{2*}\right)^2 - a_{12}\left(\kappa^{2*}\right)^2,\\
\Delta'_{22} &= a_{12} + a_{11}\kappa^{2*} - 2a_{12}\kappa^{2*} - a_{11}\left(\kappa^{2*}\right)^2 + a_{12}\left(\kappa^{2*}\right)^2.
\end{aligned} \tag{3.24}$$

The steady state probability of the RPD selecting action 1 is given by

$$\kappa_1 = \kappa_{11} + \kappa_{12}. \tag{3.25}$$

We have the following theorem.

Theorem 3.1. *Collectives (C4) and (C5) converge to the mixed optimal strategy, that is,*

$$\lim_{N\to\infty} \kappa_1 = \kappa^{1*} \tag{3.26}$$

if and only if $m_{12} = m_{21}$.

Hence, when the payoff matrix is symmetric, the RPDs can find their mixed optimal strategies if N is large enough. Figures 3.5 and 3.6 illustrate Theorem 3.1 for the nonsymmetric payoff matrix,

$$\begin{bmatrix} 0.4 & 0.2 \\ 0.1 & 0.3 \end{bmatrix}, \tag{3.27}$$

and the symmetric payoff matrix,

$$\begin{bmatrix} 0.4 & 0.2 \\ 0.2 & 0.3 \end{bmatrix}, \tag{3.28}$$

respectively.

The results presented in this section are a characterization of RPDs behavior in zero sum 2×2 matrix games.

4. Collectives of RPDs under fractional interactions

4.1. Environment and steady state probabilities. Consider a collective of M RPDs described as follows.

(i) $\mathcal{S}^j = \{x_1, x_2\}$ for $j = 1, 2, \ldots, M$.

(ii) Function ϕ^j in (2.1) satisfies

$$\phi^j\left(x^1, x^2, \ldots, x^j, \ldots, x^M\right) = \phi\left(\nu, x^j\right), \tag{4.1}$$

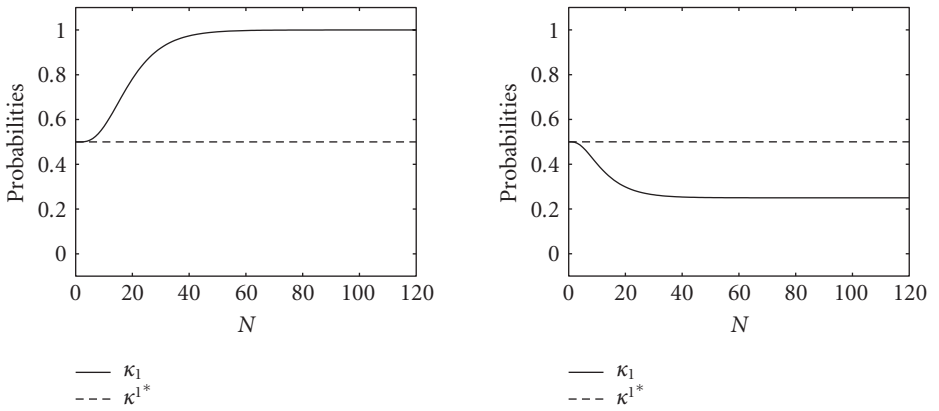

Figure 3.5. Steady state probabilities κ_1 versus N for the collectives (C4) and (C5) and payoff matrix in (3.27).

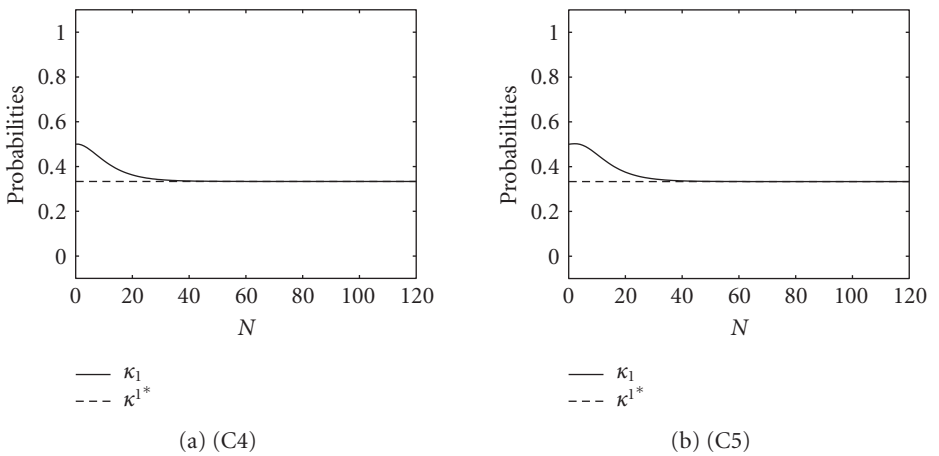

Figure 3.6. Steady state probabilities κ_1 versus N for the collectives (C4) and (C5) and payoff matrix in (3.28).

where x^i, $i = 1,2,\ldots,M$, is the state of the ith member of the collective and ν is the fraction of $x^1,x^2,\ldots,x^j,\ldots,x^M$ being in state x_1 and

$$0 < \phi(\nu, x^j) < 1. \tag{4.2}$$

(iii) Equation (3.5) or (3.6) holds if the jth RPD is L-RPD or G-RPD, respectively.

(iv) $N^j = N$, $j = 1,2,\ldots,M$, where $N \in (0,\infty)$.

(v) The RPDs make state transitions according to the sequential mode of operation.

We analyze the behavior of collectives consisting of all L-RPDs or all G-RPDs.

Let $\kappa(n) = [\kappa_1(n) \quad \kappa_2(n) \quad \cdots \quad \kappa_M(n)]$ be a row vector, where $\kappa_k(n)$ is the probability that k RPDs are in state x_1 at time n, and $\nu_k = k/M$. Then, by assumptions (i)–(v), the dynamics of the collective is described by an ergodic Markov chain,

$$\begin{aligned}
\kappa_0(n+1) &= \kappa_0(n)(1 - p(\nu_0, x_2, N)) + \kappa_1(n)(\nu_1 p(\nu_1, x_1, N)),\\
\kappa_M(n+1) &= \kappa_{M-1}(n)(\nu_1 p(\nu_{M-1}, x_2, N)) + \kappa_M(n)(1 - p(\nu_M, x_1, N)),\\
\kappa_k(n+1) &= \kappa_{k-1}(n)(\nu_{M-k+1} p(\nu_{k-1}, x_2, N))\\
&\quad + \kappa_k(n)(\nu_k(1 - p(\nu_k, x_1, N)) + \nu_{M-k}(1 - p(\nu_k, x_2, N)))\\
&\quad + \kappa_{k+1}(n)(\nu_{k+1} p(\nu_{k+1}, x_1, N)), \quad \text{for } 0 < k < M,
\end{aligned} \tag{4.3}$$

where

$$\begin{aligned}
p(\nu_k, x_i, N) &= (\phi(\nu_k, x_i))^N,\\
p(\nu_k, x_i, N) &= \begin{cases} \dfrac{(\phi(\nu_k, x_1))^N}{(\phi(\nu_k, x_1))^N + (\phi(\nu_{k-1}, x_2))^N} & \text{if } i = 1,\\[2ex] \dfrac{(\phi(\nu_k, x_2))^N}{(\phi(\nu_k, x_2))^N + (\phi(\nu_{k+1}, x_1))^N} & \text{if } i = 2, \end{cases}
\end{aligned} \tag{4.4}$$

for L-RPDs and G-RPDs, respectively.

More compactly, the dynamics can be written as

$$\kappa(n+1) = \kappa(n)A, \tag{4.5}$$

where A is a transition matrix defined by (4.3).

Let $\kappa = [\kappa_1 \quad \kappa_2 \quad \cdots \quad \kappa_M]$ be a row vector, where κ_k denotes the steady state probability of k RPDs being in state x_1. Then, (4.3) implies

$$\kappa_k = \frac{C_k^M \prod_{n=0}^{k-1} p(\nu_n, x_2, N)}{\prod_{l=1}^{k} p(\nu_l, x_1, N)} \kappa_0 \quad \forall 1 \le k \le M, \tag{4.6}$$

where

$$\kappa_0 = \frac{1}{1 + \sum_{n=1}^{M} \left(C_n^M \prod_{k=0}^{n-1} p(\nu_k, x_2, N) / \prod_{l=1}^{n} p(\nu_l, x_1, N)\right)}. \tag{4.7}$$

4.2. Homogeneous fractional interaction. In this subsection, we assume

$$\phi(\nu, x^j) = f(\nu), \quad x^j \in \{x_1, x_2\}, \tag{4.8}$$

where

$$f : [0,1] \longrightarrow (0,1), \tag{4.9}$$

is a continuous function with a unique global minimum at $\nu^* \in (0,1)$. Relationship (4.8) implies that all the RPDs have the same penalty, which depends on the fraction of the collective in state x_1. For both cases, where the collective consists of all L-RDPs and all G-RPDs, the steady state probabilities in (4.6) reduce to the same expression,

$$\kappa_k = \frac{C_k^M}{f^N(\nu_k)\sum_{l=0}^{M}\left(C_l^M/f^N(\nu_l)\right)} \quad \forall 0 \le k \le M. \tag{4.10}$$

4.2.1. Steady state behavior. The following analysis question is addressed.

A1: can the RPDs distribute themselves between x_1 and x_2 optimally, that is, so that $f(\nu)$ reaches its global minimum, $f(\nu^*)$?

Let $I = \{0,1,\ldots,M\}$ and $T \subset I$ so that for all $k \in T$,

$$f(\nu_k) = \min_{l\in I} f(\nu_l). \tag{4.11}$$

The following theorems answer this question.

Theorem 4.1. *Consider a collective of M L-RPDs or G-RPDs with homogeneous fractional interactions. Then,*

$$\lim_{a\to\infty} \sum_{k\in T} \kappa_k = 1. \tag{4.12}$$

Hence, for a collective with fixed size, the RPDs are able to distribute themselves between states x_1 and x_2 optimally if N is large enough.

Let $\nu(n)$ be the fraction of the collective in state x_1 at time n. We have the following theorem.

Theorem 4.2. *Consider a collective of L-RPDs or G-RPDs with homogeneous fractional interactions and fixed N. Moreover, assume the penalty function f is Lipschitz. Then,*

$$\lim_{M\to\infty}\lim_{n\to\infty} \nu(n) = \frac{1}{2} \quad \textit{in probability.} \tag{4.13}$$

Therefore, as the size of a collective becomes arbitrarily large while N is fixed, the RPDs distribute themselves equally between the two states. This behavior is similar to that of a statistical mechanical gas and is referred to as convergence to maximum entropy.

From Theorems 4.1 and 4.2, one can see that the parameters N and M have opposing effects. Increasing N increases the ability of the RPDs to sense the difference between the two states, while increasing M reduces this ability. Thus, we ask the following question.

A2: can the RPDs distribute themselves between states x_1 and x_2 optimally as N and M increase simultaneously?

Let Δ be a sufficiently small number and define the intervals,

$$A = \left[\nu^* - \frac{\Delta}{2}, \nu^* + \frac{\Delta}{2}\right], \tag{4.14}$$

$$B = \left[\frac{1}{2} - \frac{\Delta}{2}, \frac{1}{2} + \frac{\Delta}{2}\right]. \tag{4.15}$$

Moreover, let $I_A \subset I$ and $I_B \subset I$ so that for all $k \in I_A$ and for all $k \in I_B$, we have $\nu_k \in A$ and $\nu_k \in B$, respectively. We have the following theorems.

Theorem 4.3. *Consider a collective of L-RPDs or G-RPDs with homogeneous fractional interactions. Given interval A, there exists a constant C_A such that if*

$$\lim_{\substack{N\to\infty \\ M\to\infty}} \frac{N}{M} > C_A, \tag{4.16}$$

one has

$$\lim_{\substack{N\to\infty \\ M\to\infty}} \frac{\sum_{k\in I_A} \kappa_k}{\sum_{k\notin I_A} \kappa_k} = \infty. \tag{4.17}$$

Hence, when both N and M grow without bound, N must grow fast enough so that (4.16) holds in order for the RPDs to distribute themselves among x_1 and x_2 optimally with high probability.

Theorem 4.4. *Consider a collective of L-RPDs or G-RPDs with homogeneous fractional interactions. Given interval B, there exists a constant C_B such that if*

$$\lim_{\substack{N\to\infty \\ M\to\infty}} \frac{N}{M} < C_B, \tag{4.18}$$

one has

$$\lim_{\substack{N\to\infty \\ M\to\infty}} \frac{\sum_{k\in I_B} \kappa_k}{\sum_{k\notin I_B} \kappa_k} = \infty. \tag{4.19}$$

Therefore, when both N and M grow without bound and M is growing so fast that (4.18) holds, the convergence to maximum entropy will take place.

4.2.2. Transient behavior. Next, the convergence rates of the collectives are analyzed. Assuming $M = 10$ and f in (4.9) is given by

$$f(\nu) = \frac{80}{49}(\nu - 0.3)^2 + 0.1, \tag{4.20}$$

Figure 4.1 shows the second largest eigenvalue, λ_2, of A in (4.5) as a function of N for collectives with all L-RPDs and all G-RPDs. Clearly, as N becomes large, it will take an arbitrary long time for L-RPDs to converge while this is not true for G-RPDs.

4.3. Nonhomogeneous fractional interaction. In this subsection, we assume

$$\phi(\nu, x^j) = \begin{cases} f_1(\nu) & \text{if } x^j = x_1, \\ f_2(\nu) & \text{if } x^j = x_2, \end{cases} \tag{4.21}$$

where

$$f_1 : [0,1] \longrightarrow (0,1) \tag{4.22}$$

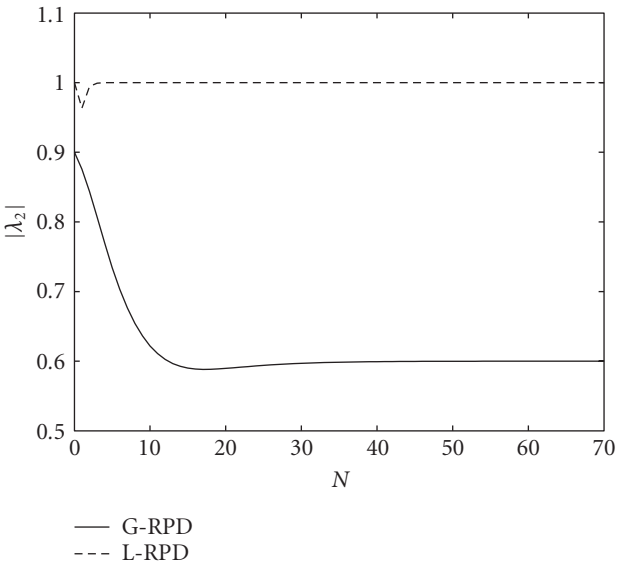

Figure 4.1. Second largest eigenvalue $|\lambda_2|$ versus N for collectives consisting of all L-RPDs and all G-RPDs.

is a continuous strictly increasing function,

$$f_2 : [0,1] \longrightarrow (0,1) \tag{4.23}$$

is a continuous strictly decreasing function, and f_1 and f_2 intersect at a single point $\nu^* \in (0,1)$. Thus, RPDs in different states are penalized differently while RPDs in the same states have the same penalties, which depend on the fraction of the collective in state x_1. Moreover, when the fraction of the RPDs in x_1 is ν^*, no RPD can decrease its penalty by changing its state while others stay in their states. Hence, ν^* is a Nash equilibrium [7]. For both cases, where the RPDs are all L-RPDs and all G-RPDs, the steady state probabilities in (4.6) become

$$\kappa_k = \frac{C_k^M \prod_{n=0}^{k-1} f_2^N(\nu_n)}{\prod_{l=1}^{k} f_1^N(\nu_l)} \kappa_0 \quad \forall\, 1 \le k \le M, \tag{4.24}$$

where

$$\kappa_0 = \frac{1}{1 + \sum_{n=1}^{M} \left(C_n^M \prod_{k=0}^{n-1} f_2^N(\nu_k) / \prod_{l=1}^{n} f_1^N(\nu_l) \right)}. \tag{4.25}$$

4.3.1. Steady state behavior. The following analysis question is addressed.

A1: can the RPDs distribute themselves between x_1 and x_2 optimally, that is, in the neighborhood of ν^*?

Let k^* be the smallest number in $\{1,2,\ldots,M\}$ such that

$$f_1(\nu_{k^*+1}) \geq f_2(\nu_{k^*}). \tag{4.26}$$

We have the following theorem.

Theorem 4.5. *Consider a collective of M L-RPDs or G-RPDs with nonhomogeneous fractional interaction. If $f_1(\nu_{k^*+1}) > f_2(\nu_{k^*})$, then*

$$\lim_{N\to\infty} \kappa_{k^*} = 1. \tag{4.27}$$

If $f_1(\nu_{k^*+1}) = f_2(\nu_{k^*})$, then

$$\lim_{N\to\infty} \kappa_{k^*+1} = \frac{M-k^*}{M+1}, \qquad \lim_{N\to\infty} \kappa_{k^*} = \frac{k^*+1}{M+1}. \tag{4.28}$$

In other words, this theorem states that if $f_1(\nu_{k^*+1}) > f_2(\nu_{k^*})$, then, when the fraction of the RPDs in x_1 is ν_{k^*}, none of the RPDs can decrease its penalty by changing its state while others stay in their states. Convergence to ν_{k^*} is optimal. Similarly, if $f_1(\nu_{k^*+1}) = f_2(\nu_{k^*})$, then, when the fraction of the RPDs in x_1 is ν_{k^*} or ν_{k^*+1}, none of the RPDs can decrease its penalty by changing its state while others stay in their states. Convergence to ν_{k^*} or ν_{k^*+1} is optimal. Hence, for a collective of fixed size and N sufficiently large, the RPDs can find an optimal distribution. Furthermore, $|\nu^* - \nu_{k^*}| \leq 1/M$. So, if M is large, the distribution is close to ν^* if N is sufficiently large.

Let Δ be a sufficiently small number and define the interval,

$$D = \left[\nu^* - \frac{\Delta}{2}, \nu^* + \frac{\Delta}{2}\right]. \tag{4.29}$$

Moreover, let $I_D \subset I$ so that for all $k \in I_D$, we have $\nu_k \in D$. We have the following theorem.

Theorem 4.6. *Consider a collective of L-RPDs or G-RPDs with nonhomogeneous fractional interactions. Given the interval D, there exists a constant C_D so that if*

$$N > C_D, \tag{4.30}$$

one has

$$\lim_{M\to\infty} \frac{\sum_{k\in I_D} \kappa_k}{\sum_{k\notin I_D} \kappa_k} = \infty. \tag{4.31}$$

Therefore, as long as N is large enough so that (4.30) is satisfied, the RPDs, unlike under homogeneous fractional interactions, do distribute themselves close to the Nash equilibrium even if the size of the collective is growing without bound.

4.3.2. Transient behavior. Next, the convergence rates of the collectives are analyzed. Assuming $M = 10$ and f_1 and f_2 in (4.22) and (4.23) are given by

$$\begin{aligned} f_1(\nu) &= 0.8\nu + 0.1, \\ f_2(\nu) &= -\frac{12}{35}\nu + \frac{31}{70}, \end{aligned} \tag{4.32}$$

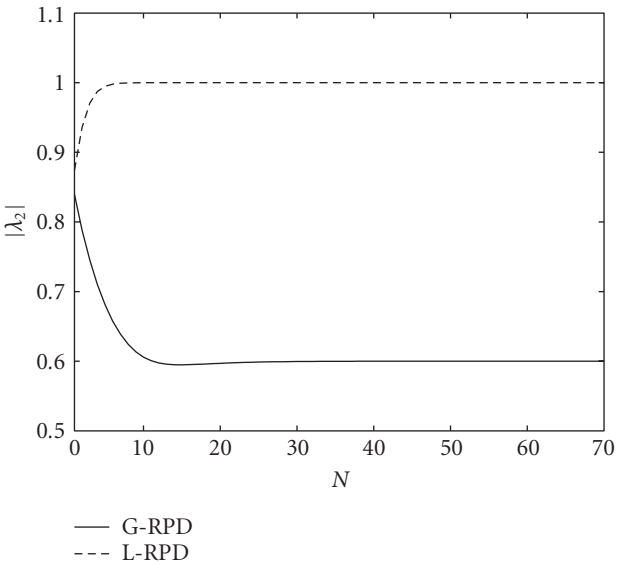

Figure 4.2. Second largest eigenvalue $|\lambda_2|$ versus N for collectives consisting of all L-RPDs and all G-RPDs.

respectively, Figure 4.2 shows the second largest eigenvalue, λ_2, of A in (4.5) as a function of N for collectives consisting of all L-RPDs and all G-RPDs. Similar to collectives with homogeneous fractional interactions, as N becomes large, it will take an arbitrarily long time for L-RPDs to converge while this is not true for G-RPDs.

4.4. Discussion. The theory presented above may be used for designing autonomously reconfigurable systems. To illustrate this, consider a robot with two operating modes, it can either perform the required work or assemble new robots. The robot is referred to as a worker or a self-reproducer when it is performing work or assembling other robots, respectively. Suppose we have a colony of such robots. For the colony to operate efficiently, there must be a right ratio between the workers and self-reproducers, depending on the environment.

To use the above theory to maintain this ratio, associate each robot with an RPD as a controller, where the states of the RPD correspond to the two operating modes of the robot. Assume the interactions of the robots are modeled as homogenous fractional with the penalty function f defined by the allocation of the robots between the worker and self-reproducer castes. The above theory suggests how the relation of the level of rationality and the size of population should be in order for the colony to sustain itself. Specifically, if the robots are not rational enough as the population becomes large, then the colony will fail to optimally distribute their operating modes. However, if the interactions of the robots are modeled as nonhomogeneous fractional with functions f_1 and f_2, as long as the level of rationality of the robots are sufficiently large, the colony will still perform optimally even if the size of the colony becomes large.

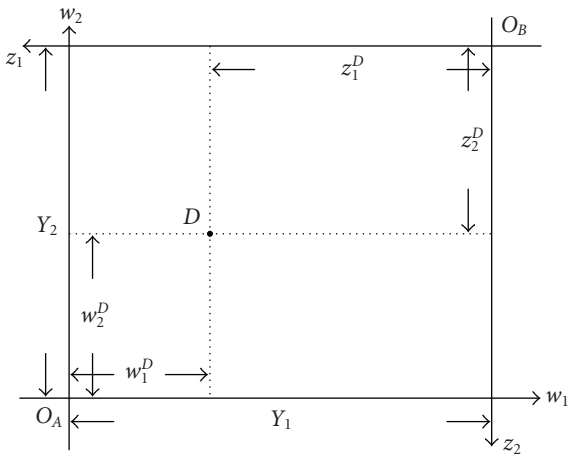

Figure 5.1. Edgeworth box describing the exchange economy.

5. Collective behavior of RPDs in Edgeworth exchange economy

5.1. Edgeworth exchange economy. Following [6], consider an exchange economy with two individuals, A and B, and two products, P_1 and P_2. The total amount of the products, P_1 and P_2, are fixed at Y_1 and Y_2 units, respectively. The allocation of the products between the two individuals exhausts the amounts of the products and is traditionally described by the Edgeworth box [6] shown in Figure 5.1. Any point in the Edgeworth box specifies a certain allocation of the products between the individuals. For example, the point D in Figure 5.1 specifies that individual A has w_1^D and w_2^D units of P_1 and P_2, respectively, and B has z_1^D and z_2^D units of P_1 and P_2, respectively, so that

$$w_1^D + z_1^D = Y_1, \qquad w_2^D + z_2^D = Y_2. \tag{5.1}$$

Note that the coordinates of D can be specified by $O_A(w_1^D, w_2^D)$ (with respect to the coordinate system with origin O_A) or $O_B(z_1^D, z_2^D)$ (with respect to the coordinate system with origin O_B).

The satisfaction of the individuals with a given allocation of the products is measured by two penalty functions. Suppose the allocation of the products between individuals A and B is at $O_A(w_1, w_2)$. Then, the penalties incurred by A and B are

$$V_A(w_1, w_2), \qquad V_B(Y_1 - w_1, Y_2 - w_2), \tag{5.2}$$

respectively, where

$$V_A : [0, Y_1] \times [0, Y_2] \longrightarrow (0, \infty), \qquad V_B : [0, Y_1] \times [0, Y_2] \longrightarrow (0, \infty). \tag{5.3}$$

The larger the penalty incurred, the less satisfied the individual.

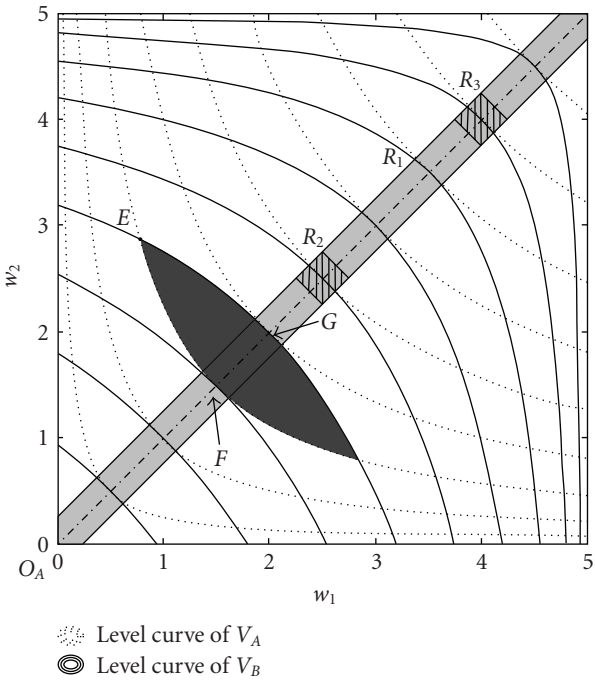

Figure 5.2. Edgeworth box showing level curves of V_A and V_B defined in (5.4).

Remark 5.1. In the economics literature, the satisfactions of the individuals are specified by utility functions [6]. The larger the value of the utility function, the greater the satisfaction of the individual. The penalty functions described above can be obtained from these utility functions, for example, by taking the reciprocal.

Figure 5.2 gives an example of the level curves of the penalty functions for A and B when

$$V_A(w_1, w_2) = (w_1 w_2)^{-0.5}, \qquad V_B(Y_1 - w_1, Y_2 - w_2) = [(Y_1 - w_1)(Y_2 - w_2)]^{-0.5}, \tag{5.4}$$

and $Y_1 = Y_2 = 5$. These penalty functions are reciprocals of the commonly used so-called Cobb-Douglas utility functions [8].

In Figure 5.2, the dash-dotted diagonal straight line, which consists of points where the level curves of V_A and V_B are tangent, is the Pareto line in the sense that when the allocation is on this line, neither A nor B can decrease its penalty by changing the allocations of the products without increasing the penalty of the other. Under classical assumptions of the Edgeworth exchange economy, the individuals are assumed to know their penalty functions exactly and never agree on exchanges that increase their penalty. Hence, in the classical model, if the allocation is initially, for instance, at point E in Figure 5.2, no matter

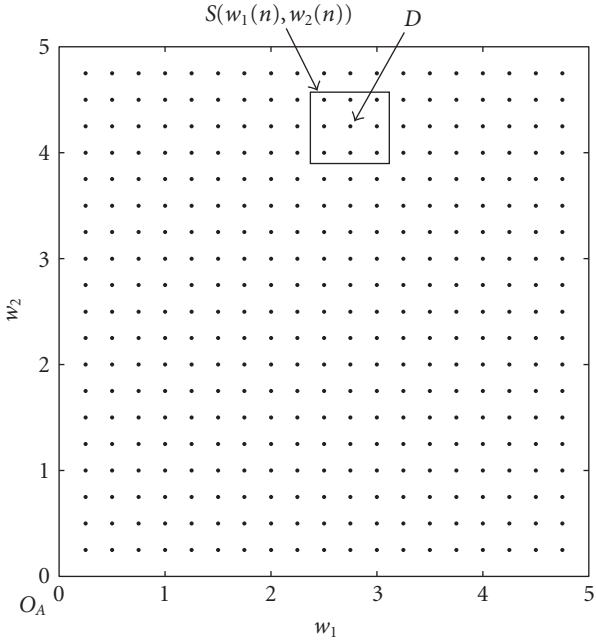

Figure 5.3. Decision space S for $Y_1 = 5$, $Y_2 = 5$, and $m = 2$.

how individuals A and B decide to exchange their products, the results of the exchanges are always in the dark grey region. Moreover, under any exchange policy, the resulting allocation will eventually converge to a point on the segment of the Pareto line, denoted in Figure 5.2 as $\overline{FG}$. When the allocation is on the Pareto line, it cannot change anymore since there will be no agreement on any exchange [6]. (The roles of sets R_1–R_3 in Figure 5.2 will become clear in Section 5.4.)

5.2. Collective of RPDs in Edgeworth exchange economy. In order to introduce RPDs in the Edgeworth exchange economy, it is convenient to discretize the decision space. Namely, let

$$\Delta = \frac{1}{2^m}, \tag{5.5}$$

where $m \in \mathbf{N}$ is given. The individuals can only have integer multiple of Δ units of both products and must have at least Δ units of each product, that is, Δ is the unit of exchange. Then, the decision space becomes

$$S = \left\{ O_A(\Delta \times l, \Delta \times h) : l \in \left\{1, 2, \ldots, \frac{Y_1}{\Delta} - 1\right\},\ h \in \left\{1, 2, \ldots, \frac{Y_2}{\Delta} - 1\right\} \right\}. \tag{5.6}$$

For $Y_1 = 5$ and $Y_2 = 5$, Figure 5.3 shows S when $m = 2$.

Assume the individuals in the Edgeworth exchange economy are modeled by G-RPDs and form a collective satisfying the following.

(i) Both A and B are G-RPDs, as defined in Section 1.2, and are referred to as the first and second G-RPD, respectively.

(ii) The G-RPDs consider only exchanges that result in at most one Δ change of their possessions. To formalize this statement, assume that $O_A(w_1(n), w_2(n))$ is the allocation after n exchanges, and let

$$T(w_1(n), w_2(n)) := \{O_A(w_1(n) + \Delta \times l,\ w_2(n) + \Delta \times h) : l \in \{-1, 0, 1\},\ h \in \{-1, 0, 1\}\}. \tag{5.7}$$

Then, the decision space of both G-RPDs is $S(w_1(n), w_2(n)) = T(w_1(n), w_2(n)) \cap S$. (See Figure 5.3 for an example of $S(w_1(n), w_2(n))$ when $m = 2$ and the allocation is at point D.) Let the states in $S(w_1(n), w_2(n))$ be denoted by x_i^n, $i \in \{1, 2, \ldots, s^n\}$, where $x_i^n = O_A(w_{1,i}^n, w_{2,i}^n)$ and s^n is the number of states in $S(w_1(n), w_2(n))$, for example, $s^n = 9$ as long as the allocation is 2Δ away from the boundaries.

(iii) The penalties for selecting state $x_i^n \in S(w_1(n), w_2(n))$ are $V_A(w_{1,i}^n, w_{2,i}^n)$ and $V_B(Y_1 - w_{1,i}^n, Y_2 - w_{2,i}^n)$ for the first and second G-RPD, respectively.

(iv) The level of rationality of the first and second G-RPD are denoted by N_1 and N_2, respectively.

(v) The probabilities of the G-RPDs selecting the states are obtained by pairwise comparison. To be more specific, let κ_i^1 be the probability of the first G-RPD selecting state x_i^n. Then,

$$\frac{\kappa_i^1}{\kappa_j^1} = \frac{F(G(N_1, V_A(w_{1,j}^n, w_{2,j}^n)/V_A(w_{1,i}^n, w_{2,i}^n)))}{F(G(N_1, V_A(w_{1,i}^n, w_{2,i}^n)/V_A(w_{1,j}^n, w_{2,j}^n)))}, \tag{5.8}$$

where functions F and G are defined in Section 1.2 and the numerator and denominator on the right-hand side of the ratio are the probabilities of the first G-RPD favoring x_i^n and x_j^n, respectively, if there were only these two choices. Assuming functions F and G are as shown in (1.9), we have

$$\frac{\kappa_i^1}{\kappa_j^1} = \left[\frac{V_A(w_{1,j}^n, w_{2,j}^n)}{V_A(w_{1,i}^n, w_{2,i}^n)}\right]^{N_1}. \tag{5.9}$$

Since $\sum_{i=1}^{s^n} \kappa_i^1 = 1$, we have

$$\kappa_i^1 = \frac{1/(V_A(w_{1,i}^n, w_{2,i}^n))^{N_1}}{\sum_{j=1}^{s^n}\left(1/(V_A(w_{1,j}^n, w_{2,j}^n))^{N_1}\right)}. \tag{5.10}$$

Similarly, we have

$$\kappa_i^2 = \frac{1/\left(V_B\left(Y_1 - w_{1,i}^n, Y_2 - w_{2,i}^n\right)\right)^{N_2}}{\sum_{j=1}^{s^n}\left(1/\left(V_B\left(Y_1 - w_{1,j}^n, Y_2 - w_{2,j}^n\right)\right)^{N_2}\right)}, \tag{5.11}$$

where κ_i^2 is the probability of the second G-RPD selecting state x_i^n.

(vi) The $n+1$st exchange made by the G-RPDs is decided by the following rule.

(a) The first and second G-RPDs propose an exchange that results in x_i^n with probabilities κ_i^1 and κ_i^2, respectively.

(b) If the proposed exchanges agree (i.e., both G-RPDs choose the same x_i^n), the exchange is made and the allocation of the products changes accordingly. Otherwise, step (a) is repeated.

The probability that the exchange results in x_i^n is

$$\frac{1/\left(V_A\left(w_{1,i}^n, w_{2,i}^n\right)\right)^{N_1}\left(V_B\left(Y_1 - w_{1,i}^n, Y_2 - w_{2,i}^n\right)\right)^{N_2}}{\sum_{j=1}^{s^n}\left(1/\left(V_A\left(w_{1,j}^n, w_{2,j}^n\right)\right)^{N_1}\left(V_B\left(Y_1 - w_{1,j}^n, Y_2 - w_{2,j}^n\right)\right)^{N_2}\right)}. \tag{5.12}$$

5.3. Scenarios. To analyze the behavior of G-RPDs in Edgeworth exchange economy, we consider the following scenarios.

(a) The level of rationalities of the G-RPDs are identical,

$$N_1 = N_2 = N, \tag{5.13}$$

$$N\Delta = 1. \tag{5.14}$$

(b) The level of rationalities of the G-RPDs are as in (5.13) but

$$N\Delta^2 = 1. \tag{5.15}$$

Hence, the G-RPDs become more rational at a faster rate as Δ becomes small so that the product of the level of rationality and Δ^2 is kept at one.

(c) The level of rationalities of the G-RPDs are not identical,

$$N_1 = N = 4N_2, \tag{5.16}$$

and N satisfies (5.15), that is, the first G-RPD is four times more rational than the second G-RPD.

5.4. Steady state behavior. The following analysis question is addressed.

A: will the allocation of the products converge to the Pareto line (and to which point on the Pareto line) as $n \to \infty$?

To investigate this question, assume, for example, that there are five units of both P_1 and P_2, that is, $Y_1 = Y_2 = 5$, in the economy and the penalty functions for the G-RPDs are as shown in (5.4). We investigate the allocation of the products as $n \to \infty$, when m in (5.5) varies from 0 to 4 in each of the scenarios (a)–(c). Although the analysis can be carried out analytically using Markov chains, the number of states increases exponentially as Δ becomes small. Hence, computer simulations are employed.

Table 5.1. Simulation results for scenario (a).

	$m = 0$	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
Δ	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
N	1	2	4	8	16	32
P_{R_1}	0.2745	0.1448	0.2698	0.2937	0.3536	0.4457

Table 5.2. Simulation results for scenario (b).

	$m = 0$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
Δ	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$
N	1	4	16	64	256
P_{R_2}	0	0.0590	0.2260	0.4678	0.8544

Table 5.3. Simulation results for scenario (c).

	$m = 0$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
Δ	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$
N	1	4	16	64	256
P_{R_3}	0.0709	0.0672	0.2234	0.4615	0.8643

The results of the analysis are as follows.

(i) The data for scenario (a) are summarized in Table 5.1. The last row indicates the frequencies of the allocation of the products converging inside R_1 (see Figure 5.2), which is the region defined by $|w_1 - w_2| \leq 0.25$. Thus, if N grows linearly with the decrease of Δ, the collective does not converge reliably to the Pareto line.

(ii) The data for scenario (b) are summarized in Table 5.2. The last row indicates the frequencies of the allocation of the products converging inside R_2 (see Figure 5.2), which is the square region centered at $O_A(5/2, 5/2)$ with area 1/8. Hence, when N grows quadratically with the decrease of Δ, the product allocation converges to a Pareto optimal and is "fair" in the sense that both G-RPDs have equal amounts of each product.

(iii) The data for scenario (c) are summarized in Table 5.3. The last row indicates the frequencies of the allocation of the products converging inside R_3 (see Figure 5.2), which is the square region centered at $O_A(4, 4)$ with area 1/8. Hence, the allocation converges to the small region around $O_A(4, 4)$, which is Pareto optimal, when N is large. Thus, when $N_1 = 4N_2$, the first G-RPD takes advantage of the second one in the sense that it ends up with four times as many of both products as the second G-RPD.

5.5. Discussion. Based on the results in Section 5.4, the following observations are made.

(i) The product allocation converges to a Pareto optimal one if the level of rationality of the G-RPDs is large enough relative to the reciprocal of the unit of exchange. Furthermore, the allocation is a fair one if the G-RPDs have the same level of rationality. However, if one of the G-RPDs has a larger level of rationality, it will take advantage of the other and end up with more products than the other.

(ii) In the classical Edgeworth exchange economy, the individuals are tacitly assumed to be of infinite rationality in the sense that they know precisely their utility functions and never accept trades that decrease their utility. As a result, the system is not ergodic, that is, the steady state depends on the initial product allocation. In the RPD formulation, the rationality is bounded [9], the system is ergodic (due to "mistakes" committed by the individuals in accepting disadvantageous trades) and, therefore, the system converges to a unique equilibrium—the Pareto point where both individuals have equal amount of products, if their rationality is the same.

6. Conclusions

Paper [1] and the current paper comprise a theory of rational behavior based on rational probabilistic deciders. This theory shows that under simple assumptions, RPDs exhibit autonomous optimal reconfigurable behavior in a large variety of situations, both individually and collectively. Among unexplored topics remains the issue of learning: in the current formulation, RPDs explore the environment every time anew, without taking into account their past experiences. Incorporating learning in RPD behavior is a major topic of future research.

Appendices

A. Proofs for Section 3

Proof of Theorem 3.1. (a) For L-RPDs.

(Sufficiency). Suppose $m_{12} = m_{21}$. Then, by (3.8) and (3.23)–(3.25), the steady state probability that the L-RPD selects state 1 is

$$\kappa_1 = \frac{-\left(\frac{1-m_{22}}{2}\right)^N (1-\kappa^{2*}) - \left(\frac{1-m_{21}}{2}\right)^N \kappa^{2*}}{-\left(\frac{1-m_{12}}{2}\right)^N - \left(\frac{1-m_{22}}{2}\right)^N - \left[\left(\frac{1-m_{11}}{2}\right)^N - \left(\frac{1-m_{12}}{2}\right)^N + \left(\frac{1-m_{21}}{2}\right)^N - \left(\frac{1-m_{22}}{2}\right)^N\right]\kappa^{2*}}. \tag{A.1}$$

Equations, (3.2) and (3.19) imply that as $N \to \infty$, $\kappa_1 \to \kappa^{2*}$. Since $m_{12} = m_{21}$ implies $\kappa^{1*} = \kappa^{2*}$, sufficiency is proved.

(Necessity). Suppose $m_{12} \neq m_{21}$. Then, (3.2), (3.19), and (A.1) imply that as $N \to \infty$, $\kappa_1 \to 1$, which is not equal to κ^{1*} for any payoff matrix satisfying (3.19) and $m_{12} \neq m_{21}$. Hence, necessity is proved.

(b) For G-RPDs.

By (3.8) and (3.23)–(3.25), the steady state probability that the G-RPD selects state 1 is

$$\kappa_1 = -\frac{\Delta_1}{\Delta}, \tag{A.2}$$

where

$$\begin{aligned}\Delta = &-\frac{(1-m_{12})^N}{(1-m_{12})^N + (1-m_{22})^N} - \frac{(1-m_{22})^N}{(1-m_{12})^N + (1-m_{22})^N} \\ &- \left[\frac{(1-m_{11})^N}{(1-m_{11})^N + (1-m_{21})^N} - \frac{(1-m_{12})^N}{(1-m_{12})^N + (1-m_{22})^N}\right. \\ &\left. + \frac{(1-m_{21})^N}{(1-m_{11})^N + (1-m_{21})^N} - \frac{(1-m_{22})^N}{(1-m_{12})^N + (1-m_{22})^N}\right]\kappa^{2*},\end{aligned} \tag{A.3}$$

$$\Delta_1 = \frac{(1-m_{22})^N}{(1-m_{12})^N + (1-m_{22})^N}(1-\kappa^{2*}) + \frac{(1-m_{21})^N}{(1-m_{11})^N + (1-m_{21})^N}\kappa^{2*}. \tag{A.4}$$

Equations (3.2), (3.19), and (A.2)–(A.4) imply that as $N \to \infty$, $\kappa_1 \to \kappa^{2*}$. Furthermore, (3.20) and (3.21) imply that $\kappa^{1*} = \kappa^{2*}$ if and only if $m_{12} = m_{21}$. Thus, the theorem is proved for G-RPDs. □

B. Proofs for Section 4

Proof of Theorem 4.1. Since (4.10) holds for collectives of all L-RPDs and all G-RPDs, the following argument holds for both: by (4.10),

$$\frac{\kappa_i}{\kappa_j} = \frac{C_i^M}{C_j^M}\left(\frac{f(\nu_j)}{f(\nu_i)}\right)^N \quad \text{for } 0 \le i, j \le M. \tag{B.1}$$

Hence, if $i \in T$ and $j \notin T$, $\kappa_i/\kappa_j \to \infty$ as $N \to \infty$. This implies $\kappa_k \to 0$ as $N \to \infty$ if $k \notin T$. Thus, (4.12) is true. □

Proof of Theorem 4.2. (a) For L-RPDs.

The dynamics of the L-RPDs with homogeneous fractional interactions can be written as

$$\nu(n+1) = \nu(n) + \frac{1}{M}\zeta(n), \tag{B.2}$$

where

$$\zeta(n) = \begin{cases} 1 & \text{with probability } (1-\nu(n))\, f^N(\nu(n)) \\ -1 & \text{with probability } \nu(n) f^N(\nu(n)) \\ 0 & \text{with probability } 1 - f^N(\nu(n)). \end{cases} \tag{B.3}$$

We note that the dynamics described above are the same as those treated in [10, 11], and we follow the discussions in [11]. Consider the dynamic system described as follows:

$$\tilde{\nu}(n+1) = \tilde{\nu}(n) + \frac{1}{M}(1 - 2\tilde{\nu}(n))\, f^N(\tilde{\nu}(n)), \tag{B.4}$$

where $\tilde{\nu}(0) = \nu(0)$. We note the following.

(i) The dynamic system in (B.4) has an equilibrium at $\nu^* = 1/2$ and, moreover, this equilibrium is global asymptotically stable.

(ii) The penalty function f is Lipschitz by assumption.

(iii) The trajectories of $\nu(n)$ in (B.2) and (B.3) are bounded.

Hence, by [10, Theorem 1], for any $\delta > 0$, we can find a number M_0 such that for all $M \geq M_0$, we have,

$$\text{Prob.}\left\{ \left| \tilde{\nu}(n) - \nu(n) \right| < \delta \right\} \leq 1 - \delta, \quad n \in [0, \infty). \tag{B.5}$$

Since $\tilde{\nu}(n) \to 1/2$ as $n \to \infty$, (B.5) implies the theorem.

(b) For G-RPDs.

Note that for the same penalty function f, a collective of G-RPDs behave in the same way as a collective of L-RPDs in the steady state. Hence, Theorem 4.2 is true for G-RPDs since, by (a), it is true for L-RPDs. □

Proof of Theorem 4.3. Since (4.10) holds for collectives of all L-RPDs and all G-RPDs, the following argument holds for both: (4.10) implies

$$\kappa_k = \mathrm{C}_k^M \kappa_0 \left(\frac{f(\nu_0)}{f(\nu_k)} \right)^N. \tag{B.6}$$

Furthermore,

$$\sum_{i=0}^{M} \mathrm{C}_i^M = 2^M. \tag{B.7}$$

Let

$$\nu^1 = \arg\inf_{\nu \notin A} f(\nu) \tag{B.8}$$

and ν^* be the global minimum of f. Then, by (B.6) and (B.7), when M is large,

$$\sum_{k \in I_A} \kappa_k > \kappa_0 \left(\frac{f(\nu_0)}{f(\nu^*)} \right)^N, \qquad \sum_{k \notin I_A} \kappa_k < 2^M \kappa_0 \left(\frac{f(\nu_0)}{f(\nu^1)} \right)^N. \tag{B.9}$$

Equation (B.9) implies

$$\frac{\sum_{k\in I_A}\kappa_k}{\sum_{k\notin I_A}\kappa_k} > \frac{1}{2^M}\left(\frac{f(\nu^1)}{f(\nu^*)}\right)^N, \tag{B.10}$$

which gives

$$\begin{aligned}\ln\frac{\sum_{k\in I_A}\kappa_k}{\sum_{k\notin I_A}\kappa_k} &> N(\ln f(\nu^1)-\ln f(\nu^*))-M\ln 2\\ &= M\left(\frac{N}{M}(\ln f(\nu^1)-\ln f(\nu^*))-\ln 2\right).\end{aligned} \tag{B.11}$$

Hence, if $N/M > \ln 2/(\ln f(\nu^1)-\ln f(\nu^*))$ as $N\to\infty$ and $M\to\infty$, (4.17) holds. Let

$$C_A = \frac{\ln 2}{\ln f(\nu^1)-\ln f(\nu^*)} \tag{B.12}$$

and the theorem is proved. □

Proof of Theorem 4.4. Since (4.10) holds for collectives of all L-RPDs and all G-RPDs, the following argument holds for both: let

$$0 < c < d < \frac{1}{2}, \tag{B.13}$$

and $[cM]$ and $[dM]$ denote the integer nearest to cM and dM, respectively. Note that,

$$\sum_{i=0}^{[cM]} \mathrm{C}_i^M < [cM]\frac{M!}{[cM]!(M-[cM])!}, \qquad \sum_{i=[dM]}^{M-[dM]} \mathrm{C}_i^M > (M-2[dM])\frac{M!}{[dM]!(M-[dM])!}. \tag{B.14}$$

To simplify nations below, define

$$S_1 = [cM]\frac{M!}{[cM]!(M-[cM])!}, \qquad S_2 = (M-2[dM])\frac{M!}{[dM]!(M-[dM])!}. \tag{B.15}$$

Define $c=(1-\Delta)/2$ (i.e., $\Delta = 1-2c$) and $d=(2-\Delta)/4$, where Δ is as shown in (4.15). Let

$$\nu^2 = \arg\max_{\nu\in B} f(\nu), \tag{B.16}$$

and ν^* be the global minimum of f. Then, when M is large, (B.6), (B.14), (B.15), and the definitions of c and d imply that

$$\sum_{i\notin I_B}\kappa_i \le 2S_1\kappa_0\left(\frac{f(\nu_0)}{f(\nu^*)}\right)^N, \qquad \sum_{i\in I_B}\kappa_i \ge S_2\kappa_0\left(\frac{f(\nu_0)}{f(\nu^2)}\right)^N. \tag{B.17}$$

Hence,

$$\frac{\sum_{i\in I_B}\kappa_i}{\sum_{i\notin I_B}\kappa_i} \geq \frac{S_2}{2S_1}\left(\frac{f(\nu^*)}{f(\nu^2)}\right)^N. \tag{B.18}$$

Note that

$$\begin{aligned}\frac{S_2}{2S_1} &= \frac{M-[dM]+1}{[cM]+1}\times\frac{M-[dM]+2}{[cM]+2}\times\cdots\times\frac{M-[cM]}{[dM]}\\ &\quad\times\frac{1-2d}{2c} \geq \lambda^{([dM]-[cM])}\frac{1-2d}{2c},\end{aligned} \tag{B.19}$$

where $\lambda=\min\{(M-[dM]+1)/([cM]+1),(M-[dM]+2)/([cM]+2),\ldots,(M-[cM])/[dM]\}$. Thus, as M becomes large, (B.18) and (B.19) imply

$$\begin{aligned}\ln\frac{\sum_{i\in I_B}\kappa_i}{\sum_{i\notin I_B}\kappa_i} &\geq ([dM]-[cM])\ln\lambda+\ln\frac{1-2d}{2c}+N(\ln f(\nu^*)-\ln f(\nu^2))\\ &\approx M\left[(d-c)\ln\lambda+\frac{N}{M}(\ln f(\nu^*)-\ln f(\nu^2))\right]+\ln\frac{1-2d}{2c}.\end{aligned} \tag{B.20}$$

Hence, if $N/M<(d-c)\ln\lambda/(\ln f(\nu^2)-\ln f(\nu^*))$ as $N\to\infty$ and $M\to\infty$, (4.19) holds. Let

$$C_B = \frac{(d-c)\ln\lambda}{\ln f(\nu^2)-\ln f(\nu^*)} \tag{B.21}$$

and the theorem is proved. □

Proof of Theorem 4.5. Since (4.24) and (4.25) hold for collectives of all L-RPDs and all G-RPDs, the following argument holds for both: by (4.24),

$$\frac{\kappa_{k+1}}{\kappa_k} = \frac{M-k}{k+1}\left(\frac{f_2(\nu_k)}{f_1(\nu_{k+1})}\right)^N \quad \text{for } 0\leq k\leq M-1. \tag{B.22}$$

(a) Suppose $f_1(\nu_{k^*+1})>f_2(\nu_{k^*})$. Then, (B.22) implies that as $N\to\infty$,

$$\frac{\kappa_{k+1}}{\kappa_k} \longrightarrow \begin{cases}\infty & \text{if } k<k^*,\\ 0 & \text{if } k\geq k^*.\end{cases} \tag{B.23}$$

Equation (B.23), the properties of f_1 and f_2, and the definition of k^* imply that, as $N\to\infty$, $\kappa_k\to 0$ for $k\neq k^*$. Hence, (4.27) is true.

(b) Suppose $f_1(\nu_{k^*+1})=f_2(\nu_{k^*})$. Then, (B.22) implies that as $N\to\infty$,

$$\frac{\kappa_{k+1}}{\kappa_k} \longrightarrow \begin{cases}\infty & \text{if } k<k^*,\\ 0 & \text{if } k>k^*.\end{cases} \tag{B.24}$$

Equation (B.24) the properties of f_1 and f_2, and the definition of k^* imply that as $N \to \infty$, $\kappa_k \to 0$ for $0 \le k < k^*$ and $k^* + 1 < k \le M$, that is,

$$\kappa_{k^*} + \kappa_{k^*+1} = 1 \quad \text{as } N \longrightarrow \infty. \tag{B.25}$$

Furthermore, $f_1(\nu_{k^*+1}) = f_2(\nu_{k^*})$ and (B.22) imply

$$\frac{\kappa_{k^*+1}}{\kappa_{k^*}} = \frac{M - k^*}{k^* + 1} \quad \forall N. \tag{B.26}$$

Equations (B.25) and (B.26) imply (4.28).

□

Proof of Theorem 4.6. Since (4.24) and (4.25) hold for collectives of all L-RPDs and all G-RPDs, the following argument holds for both: let

$$\psi(\nu_k) = \sum_{i=1}^{\nu_k M} [\ln f_2^N(\nu_{i-1}) - \ln f_1^N(\nu_i)]. \tag{B.27}$$

Then, by (4.24),

$$\kappa_k = \kappa_0 \mathrm{C}_k^M \exp(\psi(\nu_k)). \tag{B.28}$$

Furthermore,

$$\begin{aligned} \frac{1}{M}\psi(\nu_k) &= \frac{1}{M}\sum_{i=1}^{\nu_k M} [\ln f_2^N(\nu_{i-1}) - \ln f_1^N(\nu_i)] \\ &= \frac{1}{M}\sum_{i=1}^{\nu_k M} \left[\ln f_2^N\left(\nu_i - \frac{1}{M}\right) - \ln f_1^N(\nu_i)\right]. \end{aligned} \tag{B.29}$$

Hence, when M is large,

$$\frac{1}{M}\psi(\nu_k) = \int_0^{\nu_k} [\ln f_2^N(\zeta) - \ln f_1^N(\zeta)]d\zeta + \rho\left(\frac{1}{M}\right), \tag{B.30}$$

or

$$\begin{aligned} \psi(\nu_k) &= M\int_0^{\nu_k} [\ln f_2^N(\zeta) - \ln f_1^N(\zeta)]d\zeta + M\rho\left(\frac{1}{M}\right), \\ &= MN\int_0^{\nu_k} [\ln f_2(\zeta) - \ln f_1(\zeta)]d\zeta + M\rho\left(\frac{1}{M}\right), \end{aligned} \tag{B.31}$$

where $\rho(1/M)$ is an error term or order $1/M$. Let

$$\hat{\psi}(\nu) = MN\int_0^{\nu} [\ln f_2(\zeta) - \ln f_1(\zeta)]d\zeta. \tag{B.32}$$

Then, by the properties of f_1 and f_2 defined in Section 4.3, the global maximum of $\hat{\psi}(\nu)$ happens at the intersection of f_1 and f_2, which is at ν^*. Moreover, define

$$\nu^1 = \arg\sup_{\nu \notin D} \hat{\psi}(\nu). \tag{B.33}$$

Then, as M becomes large, (B.7), (B.28), (B.31), and (B.32) imply

$$\sum_{k \notin I_D} \kappa_k \le \kappa_0 2^M \exp\left(\hat{\psi}(\nu^1) + M\rho\left(\frac{1}{M}\right)\right), \qquad \sum_{k \in I_D} \kappa_k \ge \kappa_0 \exp\left(\hat{\psi}(\nu^*) + M\rho\left(\frac{1}{M}\right)\right). \tag{B.34}$$

Hence,

$$\frac{\sum_{k \in I_D} \kappa_k}{\sum_{k \notin I_D} \kappa_k} \ge 2^{-M} \exp\left(\hat{\psi}(\nu^*) - \hat{\psi}(\nu^1) + M\rho\left(\frac{1}{M}\right)\right). \tag{B.35}$$

Note that, by the Mean Value theorem,

$$\begin{aligned} \hat{\psi}(\nu^*) - \hat{\psi}(\nu^1) &= MN \int_{\nu^1}^{\nu^*} [\ln f_2(\zeta) - \ln f_1(\zeta)] d\zeta \\ &= MN(\nu^* - \nu^1)[\ln f_2(\nu_0) - \ln f_1(\nu_0)], \end{aligned} \tag{B.36}$$

where ν_0 is in between ν^* and ν^1. Equations (B.35) and (B.36) imply that as M becomes large,

$$\ln \frac{\sum_{k \in I_D} \kappa_k}{\sum_{k \notin I_D} \kappa_k} \ge M\left(-\ln 2 + N(\nu^* - \nu^1)[\ln f_2(\nu_0) - \ln f_1(\nu_0)] + \rho\left(\frac{1}{M}\right)\right). \tag{B.37}$$

Hence, as long as $N > \ln 2/(\nu^* - \nu^1)(\ln f_2(\nu_0) - \ln f_1(\nu_0))$ as $M \to \infty$, (4.31) holds. Let

$$C_D = \frac{\ln 2}{(\nu^* - \nu^1)(\ln f_2(\nu_0) - \ln f_1(\nu_0))} \tag{B.38}$$

and the theorem is proved. □

References

[1] P. T. Kabamba, W.-C. Lin, and S. M. Meerkov, "Rational probabilistic deciders: Part I. Individual behavior," to appear in *Mathematical Problems in Engineering*.

[2] V. I. Krinskiy, "Zero-sum games for two asymptotically optimal sequences of automata," *Problems of Information Transmission*, vol. 2, no. 2, pp. 43–53, 1966.

[3] V. Y. Krylov and M. L. Tsetlin, "Games between automata," *Automation and Remote Control*, vol. 24, no. 7, pp. 889–899, 1963.

[4] B. G. Pittel, "The asymptotic properties of a version of the gur game," *Problems of Information Transmission*, vol. 1, no. 3, pp. 99–112, 1965.

[5] S. M. Meerkov, "Mathematical theory of behavior—individual and collective behavior of retardable elements," *Mathematical Biosciences*, vol. 43, no. 1-2, pp. 41–106, 1979.

[6] S. E. Landsburg, *Price Theory and Applications*, South-Western College, Cincinnati, Ohio, USA, 5th edition, 2002.

[7] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951.
[8] A. Mas-Colell, M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, New York, NY, USA, 1995.
[9] H. A. Simon, *Models of Man*, John Wiley & Sons, New York, NY, USA, 1957.
[10] S. M. Meerkov, "Simplified description of slow Markov walks.I," *Automation and Remote Control*, vol. 33, no. 3, part 1, pp. 404–414, 1972.
[11] S. M. Meerkov, "Simplified description of slow Markov walks.II," *Automation and Remote Control*, vol. 33, no. 5, pp. 761–764, 1972.

P. T. Kabamba: Department of Electrical Engineering and Computer Science,
University of Michigan, Ann Arbor, MI 48109-2122, USA
Email address: kabamba@umich.edu

W.-C. Lin: Department of Electrical Engineering and Computer Science,
University of Michigan, Ann Arbor, MI 48109-2122, USA
Email address: wenchiao@umich.edu

S. M. Meerkov: Department of Electrical Engineering and Computer Science,
University of Michigan, Ann Arbor, MI 48109-2122, USA
Email address: smm@eecs.umich.edu