

Article

# Equilibrium Selection under the Bayes-Based Strategy Updating Rules

Changheng Zhao <sup>1</sup> and Jiaying Li <sup>2,\*</sup>

<sup>1</sup> School of Economics & Management, Wuhan University, Wuhan 430072, China; chzhao@whu.edu.cn

<sup>2</sup> School of Economics & Management, Hubei University of Science and Technology, Xianning 437100, China

\* Correspondence: lijiaying@whu.edu.cn

Received: 17 March 2020; Accepted: 17 April 2020; Published: 5 May 2020



**Abstract:** In this paper, first, an evolutionary game model for Bayes-based strategy updating rules was constructed, in which players can only observe a signal that reveals a strategy type instead of the strategy type directly, which deviates from the strategy type of players. Then, the equilibrium selection of populations in the case of the asymmetric game, the Battle of the Sexes (BoS), and the case of a symmetric coordination game was studied where individuals make decisions based on the signals released by each player. Finally, it was concluded that in the BoS game, when the accuracy of the signal is low, the population eventually reaches an incompatible state. If the accuracy of the signal is improved, the population finally reaches a coordinated state. In a coordination game, when the accuracy of the signal is low, the population will eventually choose a payoff-dominated equilibrium. With the improvement of signal accuracy, the equilibrium of the final selection of the population depends on its initial state.

**Keywords:** asymmetric games; evolutionary dynamics; ergodic distribution

## 1. Introduction

Because there are many Nash equilibria in a coordination game, it is difficult to predict the final game equilibrium. Individuals constantly change their strategies in the process of the game, which, as a dynamic adjustment, is a process of equilibrium selection. In evolutionary game theory, individuals with bounded rationality adjust their strategies in the process of natural selection [1–6] and finally reach equilibrium.

In a uniformly mixed infinite population [5,7], the evolutionary stability of game equilibrium was studied by constructing a deterministic evolutionary dynamic process of replicator dynamics [1,4]. Then, a stochastic replicator dynamic model [8–15] was constructed based on the determined evolutionary dynamic model and considered the impact of external random shocks to study the impact of external random disturbances on population equilibrium selection. For uniformly mixed finite populations [6], a Markov chain-based stochastic evolutionary dynamic model was constructed [16–23] to study the evolution of population strategies. For ergodic Markov processes, the stationary distribution of the population state space was calculated [17]. For non-ergodic Markov processes, the rooting probability of the absorption state of the process was calculated [16,24].

In the above studies, it is assumed that players can directly observe the strategies adopted by other players and the payoffs obtained. Players update their strategies based on the payoffs that determine the probability of choosing one of the strategies. However, in the real world, it is often difficult to directly observe the strategies adopted by opponents, which need to be considered based on the signals released by them. Therefore, some researchers have begun to use the Bayes Theorem to construct learning and updating rules based on Bayesian inference to study the mechanism of collective

behavior formation [25–30] and to analyze the equilibrium selection of the population in the long-term evolution process.

Fudenberg and Imhof (2006) [17] studied a symmetric game with a rare mutation rate in a finite population. They construct a Markov chain where the transition probabilities were related to their payoff. Their research shows that the evolutionary process simplifies significantly and approximates a simpler process over the pure states, which dramatically reduces the state space and the calculation of the stationary distribution. Based on the research of Fudenberg and Imhof (2006), Veller and Hayward (2016) [18] further studied an asymmetric game with a rare mutation rate in a finite population and presented the analytic results of the stationary distribution of the evolutionary process. However, recent studies have identified Bayesian inference as a mechanism behind collective behavior in nature and human society [26,27]. Therefore, it is necessary to analyze the equilibrium selection from the perspective of Bayesian inference.

In this paper, strategy updating rules based on Bayesian inference were constructed, in which players cannot directly observe the payoffs and strategies of their opponents, but only receive the related signals and then analyze their strategy evolution under the updating rules. For the problem of multiple equilibrium selection in a coordinated game, these rules were used to analyze the equilibrium selection in the long-term evolution process. The equilibrium choice in an asymmetric game represented by the Battle of the Sexes game (BOS) [17,18] and a symmetric game were studied, respectively. In the asymmetric Battle of the Sexes game [31–34], players with different roles use different strategies. The male player prefers a football match, while the female player prefers a ballet performance. They have to convince each other, otherwise they will get nothing. For the collective payoff, it is optimal for them to choose the same strategy. Players hold different preferences but hope to adopt the same strategy. Contrary to the Battle of the Sexes game, in a symmetric coordination game, players hold the same preferences and hope to adopt the same strategy. We constructed the evolutionary game model with strategy updating rules based on Bayesian inference and studied the evolutionary equilibrium selection in these games.

## 2. Model in the BoS Game

Suppose a  $2 \times 2$  Battle of the Sexes game has two populations, men (m) and women (w), with the size of  $N$ . At each time  $t$ , a population is randomly selected from two populations to update the strategy, and all players in the population update the strategy at the same time. For each individual in the two populations, there is the same strategy space, and the strategy type is  $S \in \{A, B\}$ . The payoff matrix of the game is as follows in Table 1:

**Table 1.** The payoff matrix of the Battle of the Sexes game.

		Women	
		A	B
Men	A	(a, b)	(0, 0)
	B	(0, 0)	(b, a)

In this matrix,  $a > b$ . There are two pure-strategy Nash equilibria  $(A, A)$  and  $(B, B)$  in this game. Both populations want to choose the same strategy. At each time  $t$ , each player cannot directly observe the strategy type of the players in the other population, but can only receive a signal  $\tilde{S}_{mt} = (\tilde{S}_{1mt}, \tilde{S}_{2mt}, \dots, \tilde{S}_{Nmt})$ ,  $\tilde{S}_{wt} = (\tilde{S}_{1wt}, \tilde{S}_{2wt}, \dots, \tilde{S}_{Nwt})$  about their strategy type. The signal of each player's strategy type is public knowledge; that is, all the players' strategy types in each population will be observed by all the players in the opposite population, the received signals of all the players' strategy types are the same. There is a deviation between the type of strategy and the signal sent by

player  $i$  and  $j$  in  $m$  and  $w$ . The conditional probability of the signal accurately reflecting the type of strategy of players  $i$  and  $j$  is as follows:

$$P(\widetilde{S}_{imt} = \widetilde{A} | S_{imt} = A) = P(\widetilde{S}_{imt} = \widetilde{B} | S_{imt} = B) = P(\widetilde{S}_{iwt} = \widetilde{A} | S_{iwt} = A) = P(\widetilde{S}_{iwt} = \widetilde{B} | S_{iwt} = B) = u \tag{1}$$

where  $1/2 < u < 1$ . Meanwhile, because every player updates their strategy according to public signal information at the same time, all players in a population will adopt the same strategy at every time  $t$ .  $P_t^m(A) = P(X(t) = N)$  and  $P_t^w(A) = P(Y(t) = N)$  represent the probability that all players in  $m$  and  $w$  adopt strategy  $A$  at time  $t$ , respectively. The strategy type of the population is inferred according to Bayesian rules updating. The posteriori probability of strategy  $A$  being selected by the opposite population considered by the population is as follows:

$$P_t^m(A | \widetilde{S}_{wt}) = \frac{u^{\delta_t^w} P_{t0}^w(A)}{u^{\delta_t^w} P_{t0}^w(A) + (1-u)^{\delta_t^w} (1 - P_{t0}^w(A))} \tag{2}$$

$$P_t^w(A | \widetilde{S}_{mt}) = \frac{u^{\delta_t^m} P_{t0}^m(A)}{u^{\delta_t^m} P_{t0}^m(A) + (1-u)^{\delta_t^m} (1 - P_{t0}^m(A))} \tag{3}$$

where  $\delta_t^w = 2n_t^w - N$ ,  $\delta_t^m = 2n_t^m - N$ ;  $n_t^w$  and  $n_t^m$  represent the number of signals of strategy  $A$  sent by the two populations  $w$  and  $m$  at time  $t$ , respectively;  $\delta_t^w$  and  $\delta_t^m$  represent the difference between the signals of strategy  $A$  and strategy  $B$  sent by the two populations at time  $t$ . As with the previous model, it is assumed in this paper that the prior probabilities of strategies  $A$  and  $B$  are equal for two populations at each time  $t$ , i.e.,  $P_{t0}^m(A) = 1/2$ ,  $P_{t0}^w(A) = 1/2$ ; the probability of choosing strategies  $A$  and  $B$  is  $1/2$  at the initial time.

In the BoS game, there is a mixed Nash equilibrium  $p^* = (p_m^*, p_w^*) = (\frac{a}{a+b}, \frac{b}{a+b})$ . For population  $m$ , if the posterior probability of population  $w$  in the state of strategy  $A$  at time  $t$  meets  $P_t^m(A | \widetilde{S}_{wt}) \geq p_w^*$ , then population  $m$  will adopt strategy  $A$ ; otherwise, population  $m$  will adopt strategy  $B$ . Similarly, for population  $w$ , if the posterior probability of the population  $m$  in the state of strategy  $A$  at time  $t$  meets  $P_t^w(A | \widetilde{S}_{mt}) \geq p_m^*$ , then population  $w$  will adopt strategy  $A$ ; otherwise, population  $w$  will adopt strategy  $B$ . From  $P_t^m(A | \widetilde{S}_{wt}) \geq p_w^*$  and  $P_t^w(A | \widetilde{S}_{mt}) \geq p_m^*$ , we derive the following:

$$\delta_t^w \log\left(\frac{1-u}{u}\right) \leq \log\left(\frac{a}{b}\right) \tag{4}$$

$$\delta_t^m \log\left(\frac{1-u}{u}\right) \leq \log\left(\frac{b}{a}\right) \tag{5}$$

Thus,  $\delta_t^w \geq \log\left(\frac{a}{b}\right) / \log\left(\frac{1-u}{u}\right)$ . Then, population  $m$  will adopt strategy  $A$ , while population  $w$  will adopt strategy  $A$  when  $\delta_t^m \geq \log\left(\frac{b}{a}\right) / \log\left(\frac{1-u}{u}\right)$ .

Because players only update the posterior probability of each other's strategy type according to the currently received signal and its prior probability at time  $t$ . At the same time, it is assumed in this paper that, at every time  $t$ , players will not be affected by the posterior probability of the previous period; that is, the prior probability of players to the other party's strategy type in each period is always equiprobable  $P_0^m(A) = P_0^w(A) = 1/2$ .  $P_t^m(A | \widetilde{S}_{wt})$  and  $P_t^w(A | \widetilde{S}_{mt})$  will only affect the current decision-making, not future ones. At this time, the strategy update of the two populations only

depends on the current signal state, and the strategy update process is a Markov process with a state space of  $\Omega = \{(A, A), (A, B), (B, A), (B, B)\}$ , and the transition probabilities are given by the following:

$$T_{(A,A) \rightarrow (A,B)} = \frac{1}{2} \left( 1 - \sum_{j=j_m^*}^N C_N^j u^j (1-u)^{N-j} \right) \tag{6}$$

$$T_{(A,A) \rightarrow (B,A)} = \frac{1}{2} \left( 1 - \sum_{j=j_w^*}^N C_N^j u^j (1-u)^{N-j} \right) \tag{7}$$

$$\begin{aligned} T_{(A,A) \rightarrow (A,A)} &= 1 - T_{(A,A) \rightarrow (A,B)} - T_{(A,A) \rightarrow (B,A)} \\ &= \frac{1}{2} \left( \sum_{j=j_m^*}^N C_N^j u^j (1-u)^{N-j} + \sum_{j=j_w^*}^N C_N^j v^j (1-u)^{N-j} \right) \end{aligned} \tag{8}$$

$$T_{(A,A) \rightarrow (B,B)} = 0 \tag{9}$$

$$T_{(A,B) \rightarrow (A,A)} = \frac{1}{2} \sum_{j=j_m^*}^N C_N^j u^j (1-u)^{N-j} \tag{10}$$

$$T_{(A,B) \rightarrow (A,B)} = 1 - \frac{1}{2} \sum_{j=j_m^*}^N C_N^j u^j (1-u)^{N-j} - \frac{1}{2} \sum_{j=j_w^*}^N C_N^j u^j (1-u)^{N-j} \tag{11}$$

$$T_{(A,B) \rightarrow (B,A)} = 0 \tag{12}$$

$$T_{(A,B) \rightarrow (B,B)} = \frac{1}{2} \sum_{j=j_m^*}^N C_N^j u^j (1-u)^{N-j} \tag{13}$$

$$T_{(B,A) \rightarrow (A,A)} = \frac{1}{2} \sum_{j=j_w^*}^N C_N^j u^j (1-u)^{N-j} \tag{14}$$

$$T_{(B,A) \rightarrow (A,B)} = 0 \tag{15}$$

$$T_{(B,A) \rightarrow (B,A)} = 1 - \frac{1}{2} \sum_{j=j_w^*}^N C_N^j u^j (1-u)^{N-j} - \frac{1}{2} \sum_{j=j_w^*}^N C_N^j u^j (1-u)^{N-j} \tag{15}$$

$$T_{(B,A) \rightarrow (B,B)} = \frac{1}{2} \sum_{j=j_w^*}^N C_N^j u^j (1-u)^{N-j} \tag{16}$$

$$T_{(B,B) \rightarrow (A,A)} = 0 \tag{17}$$

$$T_{(B,B) \rightarrow (A,B)} = \frac{1}{2} \left( 1 - \sum_{j=j_m^*}^N C_N^j u^j (1-u)^{N-j} \right) \tag{18}$$

$$T_{(B,B) \rightarrow (B,A)} = \frac{1}{2} \left( 1 - \sum_{j=j_w^*}^N C_N^j u^j (1-u)^{N-j} \right) \tag{19}$$

$$T_{(B,B) \rightarrow (B,B)} = \sum_{j=j_m^*}^N C_N^j u^j (1-u)^{N-j} + \sum_{j=j_w^*}^N C_N^j u^j (1-u)^{N-j} \tag{20}$$

where  $j_w^* = \max\left\{0, \left\lceil \left( N + \frac{\log \frac{a}{b}}{\log \left( \frac{1-u}{u} \right)} \right) / 2 \right\rceil \right\}$ ,  $j_m^* = \left\lfloor \left( N + \frac{\log \frac{b}{a}}{\log \left( \frac{1-u}{u} \right)} \right) / 2 \right\rfloor$ . Meanwhile, let  $\Phi_w = \sum_{j=j_w^*}^N C_N^j u^j (1-u)^{N-j}$ ,  $\Phi_m = \sum_{j=j_m^*}^N C_N^j u^j (1-u)^{N-j}$ . Similarly, the transition probability matrix of Markov process is as follows:

$$T = \begin{pmatrix} \frac{\Phi_m + \Phi_w}{2} & \frac{1 - \Phi_m}{2} & \frac{1 - \Phi_w}{2} & 0 \\ \frac{\Phi_m}{2} & 1 - \Phi_m & 0 & \frac{\Phi_m}{2} \\ \frac{\Phi_w}{2} & 0 & 1 - \Phi_w & \frac{\Phi_w}{2} \\ 0 & \frac{1 - \Phi_m}{2} & \frac{1 - \Phi_w}{2} & \frac{\Phi_m + \Phi_w}{2} \end{pmatrix} \quad (21)$$

**Proposition 1.** When  $\frac{a}{b} < \left(\frac{u}{1-u}\right)^N$ , the Markov process is ergodic with a stable distribution as follows:

$$\pi = (\pi_{A,A}, \pi_{A,B}, \pi_{B,A}, \pi_{B,B}) = \frac{1}{\Phi_m + \Phi_w} (\Phi_m \Phi_w, \Phi_w (1 - \Phi_m), \Phi_m (1 - \Phi_w), \Phi_m \Phi_w) \quad (22)$$

According to Proposition 1, when  $t \rightarrow \infty$ , the population is in a coordinated state; that is, the probability of  $(A, A)$  and  $(B, B)$  states is equal. At the same time, it is derived from  $\Phi_w \geq \Phi_m$  that  $\pi_{A,B} > \pi_{B,A}$ . When the population is in an incompatible state, strategy A will bring higher payoffs to m, while strategy B will bring higher payoffs to population w. Compared with the incompatible state  $(B, A)$ , the population will choose the incompatible state  $(A, B)$  with higher probability.

**Proposition 2.** When  $\frac{a}{b} \geq \left(\frac{u}{1-u}\right)^N$ , population m and population w will eventually be in an incompatible state  $(A, B)$ .

It is proven that when  $\frac{a}{b} \geq \left(\frac{u}{1-u}\right)^N$ ,  $\Phi_w = 1$  and  $\Phi_m = 0$ . At this time, the matrix of transition probability of the Markov process becomes:

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \quad (23)$$

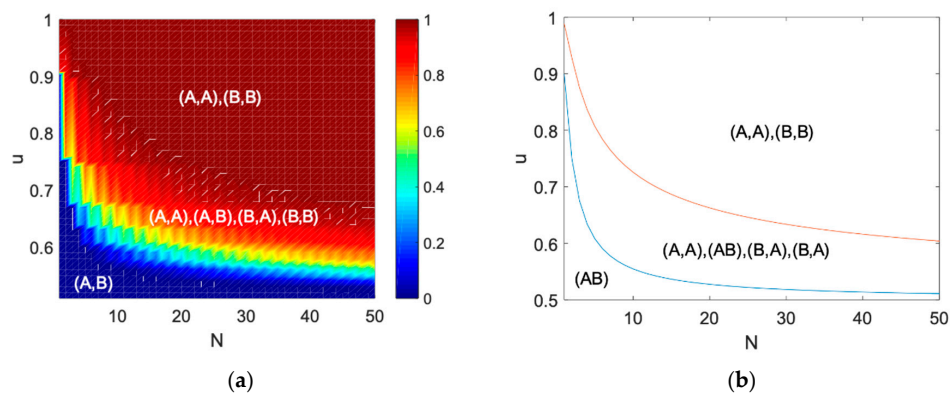
In the Markov process, states  $(A, A)$ ,  $(B, A)$ , an  $(B, B)$  are transient, and state  $(A, B)$  is the absorbing state of the process.

Proposition 2 shows that in this asymmetric game, if there is a big difference in payoff between the two kinds of coordinated equilibria, the population will choose the strategy that brings higher payoffs to itself, which makes the two populations always in an incompatible state, unable to reach the Nash equilibrium.

In Figure 1, (a) shows the probability that the population is in coordination  $(A, A)$  or  $(B, B)$  under the combination of different parameters  $u$  and  $N$ . Figure 1b shows the final evolutionarily stable state of the strategy updating process in the region with different parameter combinations.

In Figure 1, (b) also shows that when the population size  $N$  is smaller or the signal reliability  $u$  is smaller, the population will be more inclined to choose the strategy that brings higher payoffs to itself. Because the two populations have different incomes under different equilibrium tendencies, population m is more inclined to choose  $(A, A)$  equilibrium, while population w is more inclined to choose  $(B, B)$  equilibrium, which makes the two populations always in an incompatible state  $(A, B)$ . With the improvement of signal accuracy, four states coexist. When the signal accuracy is further improved, the population must be in an  $(A, A)$  or  $(B, B)$  equilibrium state after a long-time evolution. It is thus concluded that the increase of signal accuracy and population size is more conducive to population coordination.

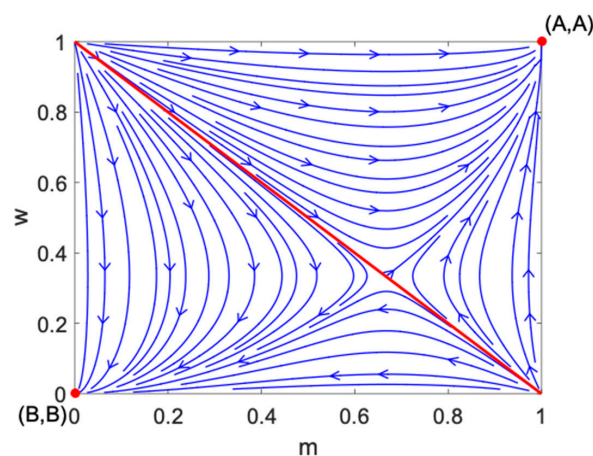
**Proposition 3.** When  $N \rightarrow \infty$ , there is a  $u^* = 1/2$ , so that when  $u > u^*$ , the population m and the population w will finally be in a coordinated state  $(A, A)$  or  $(B, B)$ .



**Figure 1.** The probability that the population reaches the coordination states (A,A) or (B,B) in the Battle of the Sexes game.

It is proven that when  $N \rightarrow \infty$ ,  $j_w^*/N \leq 1/2 < u$ , from the law of large numbers,  $\lim_{N \rightarrow \infty} \Phi_w = \lim_{N \rightarrow \infty} P(\sum_{i=1}^N \frac{X_i}{N} \geq \frac{j_w^*}{N}) = 1$ . Meanwhile, let  $f(u) = (N + \frac{\log \frac{b}{a}}{\log(\frac{1-u}{u})})/2N$ , when  $u > 1/2$ ,  $f'(u) < 0$ , and  $f(1/2) \rightarrow +\infty$ ,  $f(1) \rightarrow 1/2$ . At this time, there must be a point  $u^*$  that makes  $f(u^*) = u^*$ , when  $u > u^*$ ,  $u > f(u)$ . When  $N \rightarrow \infty$ ,  $\lim_{N \rightarrow \infty} f(u) = 1/2$ , and then  $u^* = 1/2$ . Similarly, according to the law of large numbers, when  $u > u^*$ ,  $\lim_{N \rightarrow \infty} \Phi_m = \lim_{N \rightarrow \infty} P(\sum_{i=1}^N \frac{X_i}{N} \geq \frac{j_m^*}{N}) = 1$ .

It is derived from Proposition 3 that when the population size area is infinite, there are two absorbing states (A,A) and (B,B) in the Markov process. After a long time of evolution, the population will finally reach the absorbing state; that is, the population will be in an equilibrium state. Based on the theoretical model of replicator dynamics of infinite population, as shown in Figure 2, (A,A), (B,B) show the evolutionary stable equilibrium of the replicator dynamic system. After a long series of games between populations, the evolutionary stable equilibrium will be achieved, and the final evolutionary stable state of the population depends on its initial state. Replicator dynamics is a deterministic imitation process. However, under the Bayes-based strategy updating rules, when the infinite population meets the assumption that the signal accuracy is  $1/2 < u < 1$ , that is, the signal has a higher probability to reflect its real strategy type, even if the players cannot directly observe the strategy type of the opponent, they can still accurately guess and update the strategy according to a large number of signals sent by the opponent's population. At this time, the population will still reach the equilibrium state.



**Figure 2.** Evolutionary stable equilibrium of the BoS game under replicator dynamics.

### 3. Model for Coordination Game

As showed in Table 2, suppose a  $2 \times 2$  coordination game  $G$  has two populations, 1 and 2 with the size of  $N$ . At each time  $t$ , a population is randomly selected from two populations to update the strategy, and all players in the population update the strategy at the same time. For each individual in the two populations, there is the same strategy space, and the strategy type is  $S \in \{A, B\}$ . The payoff matrix of the game is as follows:

**Table 2.** The payoff matrix of the coordination game.

		Player 1	
		A	B
Player 2	A	(a, a)	(0, 0)
	B	(0, 0)	(b, b)

In this matrix,  $a > b$ . There are two pure-strategy Nash equilibria  $(A, A)$  and  $(B, B)$  in this game. Both populations want to choose the same strategy. Same as the previous model, at each time  $t$ , each player cannot directly observe the strategy type of the players in the other population, but can only receive the signal  $\tilde{S} = (\tilde{S}_{1t}, \tilde{S}_{2t}, \dots, \tilde{S}_{Nt})$  about the their strategy type. The signal of each player’s strategy type is public knowledge; that is, all the players’ strategy types in each population will be observed by all the players in the opposite population, and the received signals of all the players’ strategy types are the same. There is a deviation between the type of strategy and the signal sent by player  $i$  in each population. The conditional probability of the signal accurately reflecting the type of strategy of player  $i$  is as follows:

$$P(\tilde{S}_{it} = \tilde{A} | S_{it} = A) = P(\tilde{S}_{it} = \tilde{B} | S_{it} = B) = u \tag{24}$$

In the coordination game, there is a mixed Nash equilibrium  $p^* = \frac{b}{a+b}$ . At this time, if the posterior probability of the other population in the state of strategy A at  $t$  time meets  $P_t^w(A | \tilde{S}_t) \geq p^*$ , then population  $w$  will adopt strategy A; otherwise, population  $w$  will adopt strategy B. It is derived from  $P_t^w(A | \tilde{S}_t) \geq p^*$  that when  $\delta_t \geq \log(\frac{a}{b}) / \log(\frac{1-u}{u})$ , the population will choose strategy A.  $j_A^* = \max\left\{0, \left\lceil \left(N + \frac{\log \frac{a}{b}}{\log(\frac{1-u}{u})}\right) / 2 \right\rceil\right\}$ ,  $J_B^* = \left\lfloor \left(N + \frac{\log \frac{b}{a}}{\log(\frac{1-u}{u})}\right) / 2 \right\rfloor$ . At the same time, let  $\Phi_A = \sum_{j=j_A^*}^N C_N^j u^j (1-u)^{N-j}$ ,  $\Phi_B = \sum_{j=j_B^*}^N C_N^j u^j (1-u)^{N-j}$ . Similarly, the matrix of transition probability of Markov process is as follows:

$$T = \begin{pmatrix} \frac{\Phi_A}{2} & \frac{1-\Phi_A}{2} & \frac{1-\Phi_A}{2} & 0 \\ \frac{\Phi_A}{2} & 1 - \frac{\Phi_A + \Phi_B}{2} & 0 & \frac{\Phi_B}{2} \\ \frac{\Phi_A}{2} & 0 & 1 - \frac{\Phi_A + \Phi_B}{2} & \frac{\Phi_B}{2} \\ 0 & \frac{1-\Phi_B}{2} & \frac{1-\Phi_B}{2} & \Phi_B \end{pmatrix} \tag{25}$$

**Proposition 4.** When  $\frac{a}{b} < \left(\frac{u}{1-u}\right)^N$ , the Markov process is ergodic with a stable distribution as follows:

$$\begin{aligned} \pi &= (\pi_{A,A}, \pi_{A,B}, \pi_{B,A}, \pi_{B,B}) \\ &= \frac{1}{2-\Phi_A-\Phi_B} (\Phi_A(1-\Phi_B), (1-\Phi_A)(1-\Phi_B), (1-\Phi_A)(1-\Phi_B), \Phi_B(1-\Phi_A)) \end{aligned} \tag{26}$$

It is derived from  $\Phi_A \geq \Phi_B$  that  $\pi_{A,A} > \pi_{B,B}$ , because the equilibrium point  $(A, A)$  will bring higher payoffs than  $(B, B)$ , the population will be in the equilibrium state  $(A, A)$  with a higher probability.

Strategy A will bring higher payoffs to population 1, while strategy B will bring higher payoffs to population 2; thus, the population will choose the incompatible state  $(A, B)$  with higher probability compared with the incompatible state  $(B, A)$ .

**Proposition 5.** When  $\frac{a}{b} \geq \left(\frac{u}{1-u}\right)^N$ , the two populations will eventually choose the payoff-dominant equilibrium  $(A, A)$ .

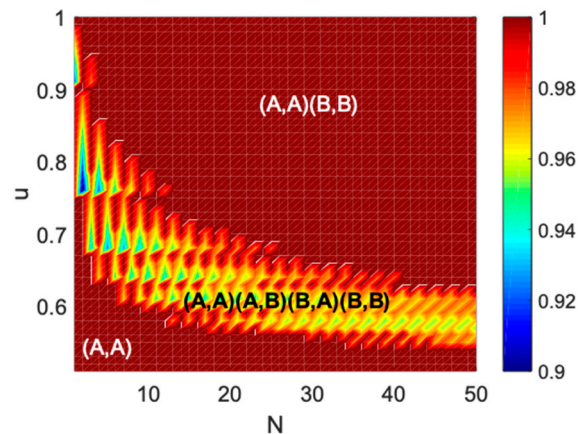
It is proved that when  $\frac{a}{b} \geq \left(\frac{u}{1-u}\right)^N$ ,  $\Phi_A = 1$  and  $\Phi_B = 0$ . At this time, the matrix of transition probability of the Markov process becomes the following:

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \quad (27)$$

In the Markov process, states  $(A, B)$ ,  $(B, A)$ , and  $(B, B)$  are transient, and state  $(A, A)$  is the absorbing state of the process.

Proposition 5 shows that in this coordination game, if there is a big difference in payoff between the two kinds of coordinated equilibria, the population will choose the equilibrium that brings higher payoffs to itself, which results in the two populations always being in a coordinated state, finally reaching the Nash equilibrium.

Figure 3 shows the probability that the population is in coordination  $(A, A)$  or  $(B, B)$  under the combination of different parameters  $u$  and  $N$ . It shows the final evolutionarily stable state of the strategy updating process in the region with different parameter combinations.



**Figure 3.** The probability that the population reaches the coordination state  $(A, A)$  or  $(B, B)$  in the coordination game.

Figure 3 shows that when the accuracy of the signal is low and the population size is small, it will be difficult for the population to determine the strategy type of the opponent through signals, and the population will prefer to choose the payoff-dominant equilibrium  $(A, A)$ . With the improvement of signal accuracy and the increase of population size, people can infer the strategy type of opponents to a certain extent. At this time, there will be the coexistence of four strategies. When the signal accuracy is very high or the population size is very large, the population can accurately determine the strategy type of the opponent and be in equilibrium. Compared with the BoS game, there is a very small chance for the probability of players in the coordination game to be in a non-equilibrium state.

**Proposition 6.** When  $N \rightarrow \infty$ , there is a  $u^* = 1/2$ , so that when  $u > u^*$ , the two populations will finally be in a coordinated state  $(A, A)$  or  $(B, B)$ .



The proof is same as that of Proposition 3. When  $N \rightarrow \infty$ ,  $\lim_{N \rightarrow \infty} \Phi_A = \lim_{N \rightarrow \infty} \Phi_B = 1$ , the matrix of transition probability at this time of this process is the following:

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{28}$$

At this time, there are two absorbing states  $(A, A)$  and  $(B, B)$ , and the final state of the process depends on its initial state.

It is derived from Proposition 6 that when the population size area is infinite, there are two absorbing states  $(A, A)$  and  $(B, B)$  in the Markov process. After a long time of evolution, the population will finally reach the absorbing state; that is, the population will be in an equilibrium state. Based on the theoretical model of replicator dynamics of infinite population, as shown in Figure 4,  $(A, A)$ ,  $(B, B)$  show the evolutionary stable equilibrium of the replicator dynamic system. After a long period of games between populations, the evolutionary stable equilibrium will be achieved, and the final evolutionary stable state of the population depends on its initial state. Replicator dynamics is a deterministic imitation process. However, under the Bayes-based strategy updating rules, when the infinite population meets the assumption that the signal accuracy is  $1/2 < u < 1$ , that is, the signal has a higher probability to reflect its real strategy type, even if the players cannot directly observe the strategy type of the opponent, they can still accurately guess and update the strategy according to a large number of signals sent by the opponent’s population. At this time, the population will still reach the equilibrium state.

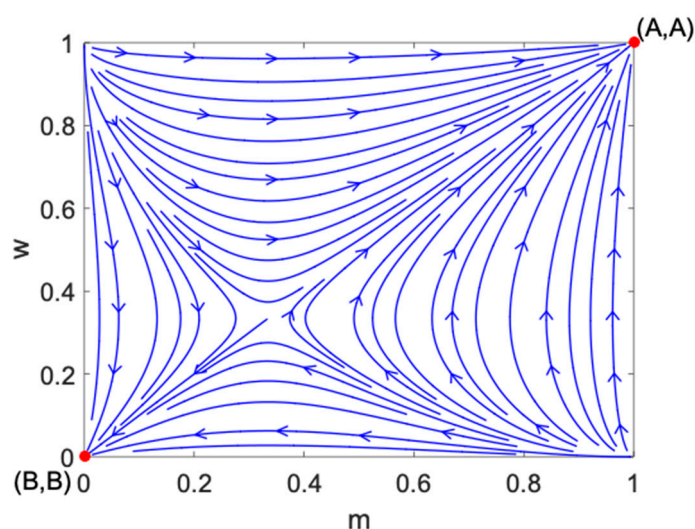


Figure 4. Evolutionary stable equilibrium of the coordination game under replicator dynamics.

#### 4. Conclusions

In this paper, we analyzed the equilibrium selection from the perspective of Bayesian inference. In the study of Veller and Hayward (2016) [18], the state of population switches among different pure states but tends to lead towards the coordination state. However, our research shows that the finite population finally reaches the incompatible state when the accuracy of the signal is low in the BoS game. With the improvement of signal accuracy, the Markov process has a stable distribution, and the population transfers between the coordinated and incompatible states. When the signal accuracy is very high, the population will eventually reach a coordinated state. Because the state of the population is closely related to the accuracy of the signal, populations always tend towards uncoordinated states because of inaccurate signal, which is different from the results of Veller and

Hayward (2016) [18]. In the infinite populations, as long as the accuracy of the signal is more than one half, the population will eventually reach a coordinated state. In the coordination game, when the accuracy of the signal is low, the population will reach the coordinated state and finally choose the payoff-dominant equilibrium, which is different from the case of the BoS game. With the improvement of signal accuracy, the population transfers between the coordinated and incompatible states and will eventually reach a coordinated state when the signal accuracy is very high, which is similar to the case of the BoS game. In the infinite populations, the population will always reach a coordinated state, and the finally selected equilibrium of the population depends on its initial state.

**Author Contributions:** C.Z. contributed to the conception of the study, J.L. performed the data analyses and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Research Foundation from Ministry of Education of China (Grant No.15JZD024) 2020 MOE Layout Foundation of Humanities and Social Sciences (Grant No.932).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Maynard, S.J. *Evolution and the Theory of Games*; USA Cambridge University Press: New York, NY, USA, 1982.
2. Axelrod, R. *The Evolution of Cooperation*; Basic Books: New York, NY, USA, 1981.
3. Weibull, J.W. *Evolutionary Game Theory*; MIT press: Cambridge, MA, USA.
4. Hofbauer, J.; Sigmund, K. *Evolutionary Games and Population Dynamics*; Cambridge university press: Cambridge, UK, 1998.
5. Hauert, C.; De, M.S.; Hofbauer, J.; Sigmund, K. Replicator Dynamics for Optional Public Good Games. *J. Theor. Biol.* **2002**, *218*, 187–194. [[CrossRef](#)] [[PubMed](#)]
6. Nowak, M.A. *Evolutionary Dynamics*; MIT press: Cambridge, MA, USA, 2006.
7. Hauert, C.; Traulsen, A.; Brandt, H.; Nowak, M.A.; Sigmund, K. Public Goods with Punishment and Abstaining in Finite and Infinite Populations. *Biol. Theory* **2008**, *3*, 114–122. [[CrossRef](#)] [[PubMed](#)]
8. Foster, D.; Young, P. Stochastic evolutionary game dynamics. *Theor. Popul. Biol.* **1990**, *38*, 219–232. [[CrossRef](#)]
9. Fudenberg, D.; Harris, C. Evolutionary dynamics with aggregate shocks. *J. Econ. Theory* **1992**, *57*, 420–441. [[CrossRef](#)]
10. Cabrales, A. Stochastic Replicator Dynamics. *Int. Econ. Rev.* **2000**, *41*, 451–481. [[CrossRef](#)]
11. Imhof, L.A. The Long-Run Behavior of the Stochastic Replicator Dynamics. *Ann. Appl. Probab.* **2005**, *15*, 1019–1045. [[CrossRef](#)]
12. Hofbauer, J.; Imhof, L.A. Time averages, recurrence and transience in the stochastic replicator dynamics. *Ann. Appl. Probab.* **2009**, *19*, 1347–1368. [[CrossRef](#)]
13. Cutili, H.G. Important Notes for a Fuzzy Boundary Value Problem. *Appl. Math. Nonlinear Sci.* **2019**, *4*, 305–314. [[CrossRef](#)]
14. Josheski, D.; Karamazova, E.; Apostolov, M. Shapley-Folkman-Lyapunov theorem and Asymmetric First price auctions. *Appl. Math. Nonlinear Sci.* **2019**, *4*, 331–350. [[CrossRef](#)]
15. Dewasurendra, M.; Vajravelu, K. On the Method of Inverse Mapping for Solutions of Coupled Systems of Nonlinear Differential Equations Arising in Nanofluid Flow, Heat and Mass Transfer. *Appl. Math. Nonlinear Sci.* **2018**, *3*, 1–14. [[CrossRef](#)]
16. Lessard, S. Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory. *Theor. Popul. Biol.* **2005**, *68*, 19–27. [[CrossRef](#)] [[PubMed](#)]
17. Fudenberg, D.; Imhof, L.A. Imitation processes with small mutations. *J. Econ. Theory* **2006**, *131*, 251–262. [[CrossRef](#)]
18. Veller, C.; Hayward, L.K. Finite-population evolution with rare mutations in asymmetric games. *J. Econ. Theory* **2016**, *162*, 93–113. [[CrossRef](#)]
19. Imhof, L.A.; Nowak, M.A. Evolutionary game dynamics in a Wright–Fisher process. *J. Math. Biol.* **2006**, *52*, 667–681. [[CrossRef](#)]
20. Wu, B.; Gokhale, C.S.; Wang, L.; Traulsen, A. How small are small mutation rates? *J. Math. Biol.* **2012**, *64*, 803–827. [[CrossRef](#)]

21. Sekiguchi, T.; Ohtsuki, H. Fixation probabilities of strategies for bimatrix games in finite populations. *Dyn. Games Appl.* **2017**, *7*, 93–111. [[CrossRef](#)]
22. Du, J.; Wu, B.; Altrock, P.M.; Wang, L. Aspiration dynamics of multi-player games in finite populations. *J. R. Soc. Interface* **2014**, *11*, 20140077. [[CrossRef](#)]
23. Nowak, M.A.; Sasaki, A.; Taylor, C.; Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **2004**, *428*, 646–650. [[CrossRef](#)]
24. Altrock, P.M.; Traulsen, A.; Nowak, M.A. Evolutionary games on cycles with strong selection. *Phys. Rev. E* **2017**, *95*, 022407. [[CrossRef](#)]
25. McNamara, J.M.; Houston, A.I. The application of statistical decision theory to animal behavior. *J. Theor. Biol.* **1980**, *85*, 673–690. [[CrossRef](#)]
26. McNamara, J.M.; Green, R.F.; Olsson, O. Bayes' theorem and its applications in animal behaviour. *Oikos* **2006**, *112*, 243–251. [[CrossRef](#)]
27. André, O. Bayesian interactions and collective dynamics of opinion: Herd behavior and mimetic contagion. *J. Econ. Behav. Organ.* **1995**, *28*, 257–274.
28. Nishi, R.; Masuda, N. Collective opinion formation model under bayesian updating and confirmation bias. *Phys. Rev. E* **2013**, *87*, 062123. [[CrossRef](#)] [[PubMed](#)]
29. Chen, H.C.; Chow, Y.Y. Equilibrium Selection in Evolutionary Games with Imperfect Monitoring. *J. Appl. Prob.* **2008**, *45*, 388–402. [[CrossRef](#)]
30. Chen, H.C.; Chow, Y.Y.; Hsieh, J. Boundedly Rational Quasi-Bayesian Learning in Coordination Games with Imperfect Monitoring. *J. Appl. Prob.* **2006**, *43*, 335–350. [[CrossRef](#)]
31. Lang, G.; Li, X. Mixed evolutionary strategies imply coexisting opinions on networks. *Phys. Rev. E* **2008**, *77*, 016108.
32. Alonso-Sanz, R. Self-organization in the spatial battle of the sexes with probabilistic updating. *Phys. A Stat. Mech. Appl.* **2011**, *390*, 2956–2967. [[CrossRef](#)]
33. Zhao, J.; Szilagyi, M.N.; Szidarovszky, F. n-person Battle of sexes games—A simulation study. *Phys. A Stat. Mech. Appl.* **2008**, *387*, 3678–3688. [[CrossRef](#)]
34. Mylius, S.D. What pair formation can do to the battle of the sexes: Towards more realistic game dynamics. *J. Theor. Biol.* **1999**, *197*, 469–485. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).