Analysis and Control of Strategic Interactions in Finite Heterogeneous Populations under Best-Response Update Rule
Ramazi, Pouria; Cao, Ming

Published in:
Proc. of 2015 IEEE 54th Annual Conference on Decision and Control (CDC)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 26-01-2020
Analysis and Control of Strategic Interactions in Finite Heterogeneous Populations under Best-Response Update Rule

Pouria Ramazi and Ming Cao

Abstract—For a finite, well-mixed population of heterogeneous agents playing evolutionary games choosing to cooperate or defect in each round of the game, we investigate, when agents update their strategies in each round using the myopic best-response rule, how the number of cooperating agents changes over time and demonstrate how to control that number by changing the agents’ payoff matrices. The agents are heterogeneous in that their payoff matrices may differ from one another; we focus on the specific case when the payoff matrices, fixed throughout the evolution, correspond to prisoner’s dilemma or snowdrift games. To carry out stability analysis, we identify the system’s absorbing states when taking the number of cooperating agents as a random variable of interest. It is proven that when all the agents update frequently enough, the reachable final states are completely determined by the available types of payoff matrices. As a further step, we show how to control the final state by changing at the beginning of the evolution, the types of the payoff matrices of a group of agents.

I. INTRODUCTION

For over three decades, biologists, economists and sociologists have utilized evolutionary game theory to understand how a certain level of cooperation can be reached among selfish interacting individuals [1][2]. Such evolutionary game theoretic models have proven to be promising in engineering network systems as well especially in the context of growing applications of networks of autonomous agents performing group tasks [3][4][5]. While simulation results have suggested there are effective mechanisms to explain the emergence and promotion of cooperation among self-interested agents [6][7], rigorous mathematical statements with proofs are still in great demand. Moreover, the existing proven mathematical results, e.g., [8][9][10], usually assume at least one of the followings: (i) an infinite population under continuous-time dynamics, (ii) deterministic dynamics, and (3) homogeneous agents. While these assumptions help to simplify the mathematical setup and gain insight into the theoretical aspects of cooperation mechanisms, an interesting research line that has not been explored to its full potential is to study a finite heterogeneous population of game-playing agents under stochastic discrete-time updating dynamics [7].

Interesting simulation results have been conducted for heterogeneous populations in [11], [12] and [13] where the agents are associated with different payoff matrices. Some mathematical statements have been provided in [14] and [15] for the myopic best response update rule [15] but only for the case when the population is homogeneous and the game is symmetric. Some researchers have studied the stochastic stability of different strategies in the population when the update rule is noisy [16][14]. Besides the results addressing the stability issue of strategically interacting populations, there is an emerging trend to investigate how to control such populations. Different methods have been suggested to control the number of agents with a specific strategy in a decision-making population. For example, in [17][18], the strategies of a group of agents are fixed to a desired strategy for the duration of the game, in [19], the payoffs of a stochastic snowdrift game are changed and in [20], some changes in terms of an emission tax rate and the price of an emission permit are applied to the payoffs of the agents.

In this paper, we continue our recent work [21] by considering a finite number of agents, each belonging to a type to which a (possibly unique) payoff matrix is associated. At each time an agent is randomly chosen from the population to update her strategy that is either to cooperate or defect, according to the myopic best response update rule. In such a population, we explicitly determine the number of cooperators of each type after a sufficiently long time, find the absorbing states, and clarify their stability properties.

Consequently, we show that based on the distribution of the types in the population, the total number of cooperators in the long run either reaches a certain steady state or fluctuates between two constants. After presenting the stability analysis, we investigate whether it is possible to control the number of cooperators in the long run by changing the types of the agents. Specifically, for a given population and types, we find how to change the types of the agents to have a desired number of cooperators in the long run.

The rest of the paper is organized as follows. In Section II, we construct the framework of our work and introduce the population game that the agents are involved in. In Section III, we provide the main stability result of this paper, and explain the evolution of the number of cooperators in each type of agent. In Section IV, we focus on controlling the total number of cooperators in the population after a sufficiently long time. It is shown how to achieve a desired reachable number of cooperators in the long run by changing the types of a group of agents.

II. POPULATION STRATEGIC DYNAMICS

We consider a finite, well-mixed population of $n$ agents that are participating in a population game evolving over time $t = 0, 1, \ldots$. Each agent can choose either to cooperate
(C) or defect (D). At each time $t$, an agent is randomly activated to update her strategy according to how well she is doing when she plays her current strategy against the average population. More specifically, the four possible payoffs of an agent $i$, $i = 1, \ldots, n$, are summarized in the $2 \times 2$ payoff matrix $A_i$:

$$A_i = \begin{pmatrix} C & D \\ R_i & S_i \end{pmatrix} T_i, P_i,$$

(1)

where the payoffs $R_i, T_i, S_i$ and $P_i$ are real numbers corresponding to strategy pairs C-against-C, D-against-C, C-against-D and D-against-D respectively. Let $s_i(t)$ denote the agent $i$’s strategy at time $t$, which is defined to be $[1 0]^T$ when agent $i$ chooses to cooperate at $t$ and defined to be $1 - [1 0]^T = [0 1]^T$ otherwise, with $1 = [1 1]^T$. Let $x_C(t)$ denote the proportion of cooperators in the whole population of $n$ agents at time $t$ and define the average population vector $s_C(t) = [x_C(t) \quad 1 - x_C(t)]^T$. Then agent $i$’s payoff at time $t$ against the average population is calculated by

$$u_i(s_i(t), s_C(t)) = s_i(t)^T A_i s_C(t).$$

The myopic best-response strategy update rule for agent $i$ dictates that agent $i$ sticks to her current strategy if her alternative strategy does not give her higher payoff, and otherwise she switches her strategy, namely

$$s_i(t + 1) = \begin{cases} s_i(t) & \text{if } u_i(s_i(t), s_C(t)) \geq u_i(1 - s_i(t), s_C(t)) \\ 1 - s_i(t) & \text{otherwise} \end{cases}.$$  

Obviously, how agent $i$ calculates her payoff, or more precisely, the exact form of her payoff matrix $A_i$ in (1), is critical for the strategy update dynamics. In this paper, we focus on those $A_i$ with special structures. We require that the entries of $A_i$ satisfy

$$T_i > R_i > \max\{S_i, P_i\}. \quad (3)$$

Payoff matrices with this property correspond to either a prisoner’s dilemma (PD) or snowdrift (SD) game. Following the setup in [21], we assume that all the $A_i$ correspond to either a specific PD game satisfying $P_i > S_i$ or correspond to possible variations of the SD game satisfying $S_i > P_i$. As it is later shown in Lemma 1, the comparison of the coefficient

$$\frac{S_i - P_i}{T_i - R_i + S_i - P_i}$$

with the ratio of cooperators determines whether agent $i$ changes her strategy when she is chosen to update. Correspondingly, all SD payoff matrices can be classified into $l > 0$ types according to their different values of the quotients

$$\frac{S_j - P_j}{T_j - R_j + S_j - P_j}$$

which have to lie strictly in (0, 1). We call all those agents with the PD payoff matrix the PD agents. We label all those $l$ types of SD payoff matrices by $1, \ldots, l$ according to the descending order of magnitude of the quotients

$$\frac{S_j - P_j}{T_j - R_j + S_j - P_j},$$

and call all those agents with the $j$th type SD payoff matrix the SD$_j$ agents.

Then there are altogether $l + 1$ types among the $n$ agents. Let $n_{PD}$ denote the number of PD agents and $n_{SD_j}, j = 1, \ldots, l$, the number of SD$_j$ agents. Then given the initial conditions of all the agents with their fixed types and initial strategies, under the best-respond strategy update rule (2), the type population defined by

$$p = (n_{SD_1}, n_{SD_2}, \ldots, n_{SD_l}, n_{PD})$$

evolves in the set

$$\mathcal{P}_n := \left\{ p \in \mathbb{Z}_+^{l+1} \mid 0 \leq p_i \leq n, \sum_{i=1}^{l+1} p_i = n \right\}.$$  

Given the types $SD_1, \ldots, SD_l$ and $PD$, the type population $p \in \mathcal{P}_n$ and under the update rule (2), denote the number of cooperators in the population at time $t$ by $n_C(t)$. We simplify this notation to $n_C(t)$ in Sec. III, when the type population $p$ is clear. Since we want to know the trends of the changes of the number of cooperators in the population, we denote the distributions of the cooperators in different types by the population state

$$a(t) := (n_{SD_1}(t), n_{SD_2}(t), \ldots, n_{SD_l}(t), n_{PD}(t)),$$

where each component of the vector is the number of cooperators in the corresponding type.

In the next section, we first study how the population state evolves over time and in particular, what its absorbing states are if there is any.

### III. Stability Analysis of the Population Game

We first describe what the absorbing states of the population state $a(t)$ look like and then prove that indeed they are the absorbing states. For the types $SD_1, \ldots, SD_l$ and $PD$ and a type population $p \in \mathcal{P}_n$, define

$$k_p := \arg \min_{k \in \mathcal{L}} \sum_{i=1}^{l} p_i \geq n_{SD_{k_p}} \quad \sum_{i=1}^{l} p_i \geq n_{SD_1} \quad l + 1$$

$$\quad \sum_{i=1}^{l} p_i < n_{SD_l},$$

where $\mathcal{L} = \{1, \ldots, l\}$, and for $j \in \mathcal{L}, n_{SD_j}$ is defined to be the value of $n_{SD_j}$ for the payoff matrix of the SD$_j$ agents. We assume that the following holds

$$[n_{SD_1}] > [n_{SD_{l+1}}] \quad \forall i \in \{1, \ldots, l - 1\}. \quad (4)$$

Define the following $(l + 1)$-dimensional row vectors

$$a^* := (n_{SD_1}, \ldots, n_{SD_{k_p-1}}, 0, \ldots, 0),$$

$$a^- := (n_{SD_1}, \ldots, n_{SD_{k_p-1}}, [n_{SD_{k_p}}] - \sum_{j=1}^{k_p-1} n_{SD_j}, 0, \ldots, 0),$$

$$a^+ := (n_{SD_1}, \ldots, n_{SD_{k_p-1}}, [n_{SD_{k_p}}] - \sum_{j=1}^{k_p-1} n_{SD_j}, 0, \ldots, 0).$$

Then define the set $\mathcal{A}$ as follows. If $k_p \neq l + 1,$

$$\mathcal{A} := \left\{ a^* \sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}} \right\} \left\{ a^-, a^+ \sum_{j=1}^{l} n_{SD_j} < n_{SD_{k_p}} \right\}.$$
and if $k_p = l+1$, then $A := \{a^*\}$. With the above definitions in hand, we are ready to present the main stability result of this section.

**Theorem 1:** Consider the types $SD_1, \ldots, SD_l$ and $PD$ and the type population $p \in \mathcal{P}_n$. When every agent is activated independently with a nontrivial probability, under the update rule (2) and for any initial population state $a(0)$, there exists some time $\tau$ such that almost surely

$$a(t) \in A \quad \forall t \geq \tau.$$  \hfill (5)

Note that the theorem implies that the set $A$ is not only invariant but also globally asymptotically attractive. According to the theorem, there exists some time $\tau$ such that for all $t \geq \tau$, it holds

$$a(t) = a^* \quad k_p \neq l+1 \text{ and } \sum_{j=1}^{k_p-1} n_{SD_j} \geq n^*_{SD_{k_p}}$$  

$$a(t) \in \{a^*, a^*_+\} \quad k_p \neq l+1 \text{ and } \sum_{j=1}^{k_p-1} n_{SD_j} < n^*_{SD_{k_p}}$$  

$$a(t) = a^* \quad k_p = l+1$$

In other words, after a finite time, almost surely the population state reaches a situation where for $i \in \{1, \ldots, k_p-1\}$ all of the $SD_i$ agents cooperate, and for $i \in \{k_p+1, \ldots, l\}$, all of the $SD_i$ agents and also all of the $PD$ agents defect. Moreover, for $k_p \neq l+1$, if $\sum_{j=1}^{k_p-1} n_{SD_j} \geq n^*_{SD_{k_p}}$, all of the $SD_{k_p}$ agents defect. If $\sum_{j=1}^{k_p-1} n_{SD_j} < n^*_{SD_{k_p}}$, then in case $a^*_+ = a^*_-$ or equivalently $n^*_{SD_{k_p}} = n^*_{SD_{k_p}^*}$, we have $n_{SD_{k_p}}^* - \sum_{j=1}^{k_p-1} n_{SD_j}$, of the $SD_{k_p}$ agents cooperate and the rest defect. However, in case $a^*_+ \neq a^*_-$, then the number of $SD_{k_p}$ agents who cooperate, belongs to the set $\{n_{SD_{k_p}}^* - \sum_{j=1}^{k_p-1} n_{SD_j}, n_{SD_{k_p}}^* - \sum_{j=1}^{k_p-1} n_{SD_{k_p}}\}$. It can be shown that the number actually fluctuates between the elements of this set, which is omitted here due to page limit.

The main idea of the proof of Theorem 1 is as follows according to a sequence of lemmas to be developed. Note that due to page limit, we omit the proofs of most of the lemmas. First the case $k_p \neq l+1$ is considered. In Lemma 2 it is shown that $A$ is an invariant set. Then through Lemmas 3 and 4 it is shown that if the number of cooperators in the population is greater or smaller than the number of cooperators indicated in $A$, it will return to the number indicated in $A$ after a finite time. In addition, with the help of Lemmas 5 and 6 it is proven that all of the $SD_1, SD_2, \ldots, SD_{k_p-1}$ agents will become cooperators after a finite number of time steps, and will not change their strategies afterwards. Similarly it can be shown that all of the $SD_{k_p+1}, SD_{k_p+2}, \ldots, SD_l$ agents will become defectors after a finite time, and will not change their strategies afterwards. Finally in view of all the proven facts, one can prove that $a$ reaches and remains in $A$ after a finite time. We skip the discussion for the case when $k_p = l+1$ due to page limit.

To carry out the proof, we need to determine the best response of each type of agents to the average population $s_C$. Define $\beta_i(s_C(t))$, the best-reply(response) for player $i$ to the average population at time $t$, as the set of pure strategies $k \in \{C, D\}$ such that no other pure strategy gives her a higher payoff against $s_C(t)$:

$$\beta_i(s_C) = \{k \in \{C, D\} : u_i(k, s_C) \geq u_i(x, s_C) \forall x \in \{C, D\}\}$$

Clearly, the best reply of each agent is completely determined by her type. We use $\beta_X(s_C(t))$ to denote the best response of an $X$-type agent to the average population at time $t$.

**Lemma 1:** For any time $t \geq 0$, the best response of a $PD$ agent is

$$\beta_{PD}(s_C(t)) = \{D\}$$

and the best response of an $SD_j$ agent, $j \in \{1, \ldots, l\}$, is

$$\beta_{SD_j}(s_C(t)) = \begin{cases} \{C\} & n_C(t) < n_{SD_j}, \\ \{C, D\} & n_C(t) = n^*_{SD_j}, \\ \{D\} & n_C(t) > n^*_{SD_j}. \end{cases}$$  \hfill (6)

**Proof:** For the $PD$ agents it is clear that the best response is always to defect according to their payoff matrix which is a $PD$ payoff matrix. For an $SD_j$ agent, in order to calculate the best reply, we have to find the pure strategy $k$ such that $u_i(k, s_C)$ is maximized. From the definition of $u_i(k, s_C)$, this is equivalent to player $i$ choosing the maximum row of the multiplication $A_i s_C$:

$$A_i s_C = A_i \begin{bmatrix} x_C \\ 1 - x_C \end{bmatrix} = \begin{bmatrix} (R_i - S_i)x_C + S_i \\ (T_i - P_i)x_C + P_i \end{bmatrix}. \hfill (7)$$

Comparing the two entries of the final right-hand vector determines the best response:

$$\beta_i(s_j) = \begin{cases} \{C\} & (R_i + P_i - T_i - S_i)x_C > P_i - S_i, \\ \{C, D\} & (R_i + P_i - T_i - S_i)x_C = P_i - S_i, \\ \{D\} & (R_i + P_i - T_i - S_i)x_C < P_i - S_i. \end{cases}$$

(8)

Now according to (3) and the fact that $A_i$ is an $SD$ payoff matrix, it holds that

$$T_i > R_i \quad \Rightarrow \quad \begin{cases} R_i - T_i + P_i - S_i < 0, \\ P_i - S_i < 0. \end{cases}$$

Hence, in view of (8), we have that

$$\beta_{SD_j}(s_C) = \begin{cases} \{C\} & x_C < \frac{S_i - P_i}{T_i - R_i + S_i - P_i}, \\ \{C, D\} & x_C = \frac{S_i - P_i}{T_i - R_i + S_i - P_i}, \\ \{D\} & x_C > \frac{S_i - P_i}{T_i - R_i + S_i - P_i}. \end{cases}$$

Multiplying the conditions in $n$ completes the proof.

As can be seen from the lemma, the best response of the $PD$ agents is always to defect. So if any of these agents is chosen to update her strategy at some time, she updates to $D$ and in case her strategy was already $D$ in the previous time step, she sticks to it. Now since each agent is chosen with a nonzero probability at all time, after some finite time, almost surely all of the $PD$ agents get the chance to update their strategies. Hence, after a finite time $t_{PD}$, almost surely (a.s.) all of the $PD$ agents become defectors. Moreover, they do not change their strategies afterwards. Hence, a.s.

$$a(t) = (n_{SD_1}^C(t), n_{SD_2}^C(t), \ldots, n_{SD_l}^C(t), 0) \quad \forall t \geq t_{PD}.$$
In the rest of this section, we consider the situation when all of the PD agents have already become defectors. That is, all time $t$ are considered to be non-less than $t_{PD}$.

In the rest of the proofs we make use of the following fact. According to the definition of $k_p$, it holds that

$$\sum_{i=1}^{j} n_{SDi} < n_{SDj}^*, \quad \forall j < k_p.$$  \hspace{1cm} (9)

Now we claim that $A$ is an invariant set under the myopic best response update rule.

Lemma 2: Let $k_p \neq 1 + 1$. If $\sum_{j=1}^{k_p-1} n_{SDj} \geq n_{SDk_p}^*$, then in case the population state vector $a(t)$ reaches $a^*$ at some time $t_r$, it will remain there afterwards, i.e.,

$$(\exists t_r : a(t_r) = a^*) \Rightarrow (a(t) = a^* \forall t \geq t_r).$$ \hspace{1cm} (10)

On the other hand, if $\sum_{j=1}^{k_p-1} n_{SDj} < n_{SDk_p}^*$, then in case the population state $a(t)$ reaches the set $\{a^r, a^+\}$ at some time $t_r$, it will remain in the set afterwards, i.e.,

$$(\exists t_r : a(t_r) \in \{a^r, a^+\}) \Rightarrow (a(t) \in \{a^r, a^+\} \forall t \geq t_r).$$ \hspace{1cm} (11)

Proof. Let $\sum_{j=1}^{k_p-1} n_{SDj} \geq n_{SDk_p}^*$ be in force. According to the definition of $a^*$, it holds that

$$n_C(t_r) = \sum_{j=1}^{k_p-1} n_{SDj} \geq n_{SDk_p}^* \geq |n_{SDk_p}^*|.$$ \hspace{1cm} (12)

Now let $\sum_{j=1}^{k_p-1} n_{SDj} < n_{SDk_p}^*$ be in force. According to the definitions of $a^-$ and $a^+$, it holds that

$$n_C(t_r) \geq \sum_{j=1}^{k_p-1} n_{SDj} + (|n_{SDk_p}^*| - \sum_{j=1}^{k_p-1} n_{SDj}) \geq |n_{SDk_p}^*|.$$ \hspace{1cm} (13)

Hence, in any case $n_C(t_r) \geq |n_{SDk_p}^*|$ holds. Hence, in view of (4),

$$n_C(t_r) > n_{SDi}^*, \forall i > k_p \overset{(6)}{=} \beta_{SDi}(s_C(t_r)) = \{D\} \forall i > k_p.$$ \hspace{1cm} (14)

Hence, if any $SD_i$ agent, $i > k_p$, is chosen at $t_r$ while her strategy is $D$ at $t_r$, she will not change her strategy at $t_r + 1$. On the other hand, all $SD_i$ agents, $i < k_p$, were already defecting at $t_r$ according to the definition of $A$. Hence,

$$n_{SDi}^*(t_r + 1) = n_{SDi}^*(t_r) = 0 \quad \forall i > k_p.$$ \hspace{1cm} (15)

Now consider the $SD_{k_p}$ agents. Let $\sum_{j=1}^{k_p-1} n_{SDj} \geq n_{SDk_p}^*$ be in force. Based on the definition of $a^*$, the following holds

$$n_C^*(t_r) = 0.$$ \hspace{1cm} (16)

On the other hand, according to (12),

$n_C(t_r) \geq n_{SDk_p}^* \Rightarrow \beta_{SDk_p}(s_C(t_r)) = \{D\}$.

Hence, due to (15),

$$n_C^*(t_r) = n_C^*(t_r) = 0.$$ \hspace{1cm} (17)

In view of (13) and (14), the above equation results in

In view of (13), we obtain

Hence, in any case $n_C(t_r) < n_{SDk_p}^*$ holds. Hence, according to (4),

$$n_C(t_r) < n_{SDi}^*, \forall i < k_p \overset{(6)}{=} \beta_{SDi}(s_C(t_r)) = \{C\} \forall i < k_p.$$ \hspace{1cm} (18)

Hence, if any $SD_i$ agent, $i < k_p$, is chosen at $t_r$ while her strategy is $C$ at $t_r$, she will not change her strategy at $t_r + 1$. On the other hand, all $SD_i$ agents, $i < k_p$, were already defecting at $t_r$ according to the definition of $A$. Hence, $n_{SDi}^*(t_r + 1) = n_{SDi}^*(t_r) = n_{SDi}, \forall i < k_p.$ \hspace{1cm} (19)

Now consider the $SD_{k_p}$ agents. Let $\sum_{j=1}^{k_p-1} n_{SDj} \geq n_{SDk_p}^*$ be in force. Based on the definition of $a^*$, the following holds

$$n_C^*(t_r) = 0.$$ \hspace{1cm} (20)

On the other hand, according to (12),

$n_C(t_r) \geq n_{SDk_p}^* \Rightarrow \beta_{SDk_p}(s_C(t_r)) = \{D\}$.

Hence, due to (15),

$$n_C^*(t_r) = n_C^*(t_r) = 0.$$ \hspace{1cm} (21)

In view of (13) and (14), the above equation results in

In view of (13), we obtain

Hence, in any case $n_C(t_r) < n_{SDk_p}^*$ holds. Hence, according to (4),

$$n_C(t_r) < n_{SDi}^*, \forall i < k_p \overset{(6)}{=} \beta_{SDi}(s_C(t_r)) = \{C\} \forall i < k_p.$$ \hspace{1cm} (22)

Hence, if any $SD_i$ agent, $i < k_p$, is chosen at $t_r$ while her strategy is $C$ at $t_r$, she will not change her strategy at $t_r + 1$. On the other hand, all $SD_i$ agents, $i < k_p$, were already defecting at $t_r$ according to the definition of $A$. Hence, $n_{SDi}^*(t_r + 1) = n_{SDi}^*(t_r) = n_{SDi}, \forall i < k_p.$ \hspace{1cm} (23)

Now consider the $SD_{k_p}$ agents. Let $\sum_{j=1}^{k_p-1} n_{SDj} \geq n_{SDk_p}^*$ be in force. Based on the definition of $a^*$, the following holds

$$n_C^*(t_r) = 0.$$ \hspace{1cm} (24)

On the other hand, according to (12),

$n_C(t_r) \geq n_{SDk_p}^* \Rightarrow \beta_{SDk_p}(s_C(t_r)) = \{D\}$.

Hence, due to (15),

$$n_C^*(t_r) = n_C^*(t_r) = 0.$$ \hspace{1cm} (25)

In view of (13) and (14), the above equation results in

In view of (13), we obtain
Hence, when an SDkp agent is chosen at \( t_r \), we have that
\[
n_{SDkp}^C(t_r + 1) = n_{SDkp}^C(t_r) + 1 - \sum_{j=1}^{k_p-1} n_{SD_j}.
\]
Moreover, in general, when the chosen agent at \( t_r \) is not an SDkp agent, we have that
\[
n_{SDkp}^C(t_r + 1) = n_{SDkp}^C(t_r).
\]
(19)

Now (16), (17), (18) and (19) imply that
\[
n_{SDkp}^C(t_r+1) \in \{ [n_{SDkp}^C] - \sum_{j=1}^{k_p-1} n_{SD_j}, [n_{SDkp}^C] - \sum_{j=1}^{k_p-1} n_{SD_j} \}.
\]
In view of (13) and (14), the above equation results in
\[
(\exists t_r: a(t_r) \in \{ a^-, a^+ \}) \Rightarrow (a(t_r + 1) \in \{ a^-, a^+ \}).
\]
Consequently, (11) can be proven by induction, which completes the proof.

The time when the number of cooperators in the population equals that in \( \mathcal{A} \), plays a role in the rest of the proof. Hence, we proceed with the following definition. For \( k_p \neq l + 1 \), define \( B \) as the set of the time steps \( t_b \) such that
\[
\begin{cases}
n(C(t_b)) = \sum_{j=1}^{k_p-1} n_{SD_j} \quad \sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SDkp}^C \quad k_p-1, \\
n(C(t_b)) \in \{ [n_{SDkp}^C], [n_{SDkp}^C] \} \quad \sum_{j=1}^{k_p-1} n_{SD_j} < n_{SDkp}^C.
\end{cases}
\]
The following two lemmas are used to show that \( B \) is infinite. In Lemma 3 it is shown that if the number of cooperators \( n_C \) exceeds the maximum number of cooperators in \( \mathcal{A} \) at some time \( t \), it will return to the number of cooperators in \( \mathcal{A} \) after some finite time. The same is shown to happen in Lemma 4 when \( n_C \) becomes less than the minimum number of cooperators in \( \mathcal{A} \).

**Lemma 3:** Let \( k_p \neq l + 1 \). If \( \sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SDkp}^C \), then in case at some time \( t \),
\[
n_C(t) > \sum_{j=1}^{k_p-1} n_{SD_j},
\]
then \( a.s. \) there exists some natural number \( t_1 \) such that
\[
n_C(t + t_1) = \sum_{j=1}^{k_p-1} n_{SD_j}.
\]
Moreover, if \( \sum_{j=1}^{k_p-1} n_{SD_j} < n_{SDkp}^C \), then in case at some time \( t \),
\[
n_C(t) > [n_{SDkp}^C],
\]
then \( a.s. \) there exists some natural number \( t_1 \) such that
\[
n_C(t + t_1) \in \{ [n_{SDkp}^C], [n_{SDkp}^C] \}.
\]

**Lemma 4:** Let \( k_p \neq l + 1 \). If \( \sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SDkp}^C \), then in case at some time \( t \),
\[
n_C(t) < \sum_{j=1}^{k_p-1} n_{SD_j},
\]
then \( a.s. \) there exists some natural number \( t_1 \) such that
\[
n_C(t + t_1) = \sum_{j=1}^{k_p-1} n_{SD_j}.
\]

Now (16), (17), (18) and (19) imply that
\[
n_C(t) \in \{ [n_{SDkp}^C], [n_{SDkp}^C] \}.
\]
Moreover, if \( \sum_{j=1}^{k_p-1} n_{SD_j} < n_{SDkp}^C \), then in case at some time \( t \),
\[
n_C(t) \in \{ [n_{SDkp}^C], [n_{SDkp}^C] \},
\]
then \( a.s. \) there exists some natural number \( t_1 \) such that
\[
n_C(t + t_1) \in \{ [n_{SDkp}^C], [n_{SDkp}^C] \}.
\]
IV. CONTROLLING THE NUMBER OF COOPERATORS

Now that we know the behavior of the population game after a sufficiently long time, we are interested in investigating whether it is possible to control the number of cooperators in the final state. While there may be several ways to achieve this goal, e.g., adding some agents to the population, fixing the strategies of some of the agents to C or D, etc., perhaps the most natural parameter to control, which makes the least changes to the structure of the population game, is the type of each agent. Changing the type of an agent can be done by modifying the parameter(s) of her payoff matrix. By changing the types of a group of agents, a new type population is acquired which leads to a possibly new number of cooperators in the final state. Note that since the log run behavior of the population state is independent from the initial strategies of the agents, it does not matter when and in what order the changes in the types are done. For sake of simplicity we assume that they all happen at time $t = 0$. So the question we attack is how to set the number of cooperators in the long run to some reference value $r$ by changing the types of a group of agents at some time step $t = 0$.

Each set of changes in the types of the agents results in some type population $p$. Moreover, if we know that type population $p$ under which the number of cooperators a.s. equals the reference value for all time greater than some constant $\tau$, we also know the type of which agent should be changed to what, by just comparing the original type population $\xi$ with $p$. For $r \in \mathbb{Z}_{\geq 0}$ and a population of size $n$, define $F(r, n)$ as the set of all feasible type-populations $p \in \mathcal{P}_n$ such that under the update rule (2), the number of cooperators a.s. equals $r$ for all time greater than some constant $\tau$, i.e.,

$$F(r, n) := \{p \in \mathcal{P}_n | \exists \tau : n_C(p, t) = r \ \forall t \geq \tau\}.$$ 

Our goal is to determine $F(r, n)$. Then it becomes clear how the types of the agents must change to have $r$ cooperators in the long run. We need the following sets for a given $n \in \mathbb{N}$, $r \in \mathbb{Z}_{\geq 0}$ and $b \in \{1, \ldots, l+1\}$:

$$F^0_1(r, n) := \{p | p \in \mathcal{P}_n, k_p = b, \sum_{i=1}^{b-1} p_i = r\},$$ 

$$F^0_2(r, n) := \{p | p \in \mathcal{P}_n, k_p = b, \sum_{i=1}^{b-1} p_i < r\}.$$ 

**Theorem 2:** Consider the types $SD_1, \ldots, SD_l$ and PD. Given $r \in \mathcal{D}_n$,

1) if there exists $b \in \{2, \ldots, l\}$ such that $n_{SD_{b-1}}^* > r > n_{SD_b}^*$, then $F(r, n) = F^0_1(r, n)$;
2) if there exists $b \in \{1, \ldots, l\}$ such that $r = n_{SD_b}^*$, then $F(r, n) = F^0_1(r, n) \cup F^0_2(r, n)$;
3) if $r \in \{0, 1, \ldots, \min \{n_{SD_1}^* - 1, n\}\}$, then $F(r, n) = F^0_1(r, n)$.

We skip the proof due to page limit.

V. CONCLUDING REMARKS

We have studied a finite heterogeneous population of game-playing agents under the myopic best response update rule. We have shown that based on the distribution of the types in the population, the total number of cooperators in the long run either reaches a certain steady state or fluctuates between two constants. Moreover, we have investigated how to control the number of cooperators, and in particular, for a given population and types, we show how to change the types of the agents to reach a desired number of cooperators in the long run.

REFERENCES