Asynchronous Decision-Making Dynamics under Best-Response Update Rule in Finite Heterogeneous Populations

Pouria Ramazi and Ming Cao

Abstract—To study how sustainable cooperation might emerge among self-interested interacting individuals, we investigate the long-run behavior of the decision-making dynamics in a finite, well-mixed population of individuals, who play collectively over time a population game. Repeatedly each individual is activated asynchronously to update her decision to either cooperate or defect according to the myopic best-response rule. The game's payoff matrices, chosen to be those of either prisoner’s dilemma or snowdrift games to underscore cooperation-centered social dilemmas, are fixed, but can be distinct for different individuals. So the overall population is heterogeneous. We first classify such heterogeneous individuals into different types according to their cooperating tendencies stipulated by their payoff matrices. Then we show that no matter what initial strategies the individuals decide to use, surprisingly one can always identify one type of individuals as a benchmark such that after a sufficiently long but finite time, individuals more cooperative compared to the benchmark always cooperate while those less cooperative compared to the benchmark defect. When such fixation takes place, the total number of cooperators in the population either becomes fixed or fluctuates at most by one. Such insight provides theoretical explanation for some complex behavior recently reported in simulation studies that highlight the puzzling effect of individuals’ heterogeneity on collective decision-making dynamics.

I. INTRODUCTION

The study on mechanisms of emergence of cooperation in self-interested social populations has attracted extensive attention in the past decades [2]–[6]. One general consensus is that human’s ability to learn plays a key role to build up sustainable collective cooperation in a competitive environment. Recent experimental studies [7]–[9] indicate that in a social group how human subjects learn to update their strategies may affect dramatically the outcome of the group’s interactions. For example, the cooperation level of a group of individuals is higher when each individual focuses more on the frequencies of the behaviors of the peers (frequency-based learning) instead of the success among the peers (success-based learning) [7]. However, how the frequency-based individuals learn remains an open question, which requires much more in-depth theoretical investigation on possible learning rules and the resulted population decision-making dynamics. Evolutionary game theory [10]–[15] has provided promising theoretical tools, and in fact postulated a variety of dynamical models under frequency-based learning [16]–[20]. In one popular model, decision makers always follow the myopic best-response update rule to maximize their utilities against their opponents based on the frequencies of their opponents’ strategies [21]–[24].

Various simulation based studies have disclosed several features of such best-response evolutionary dynamics. For homogeneous populations where individuals have the same tendency to cooperate, people have identified the surprising suppression of network effects in different game setups [25]–[27]. For heterogeneous populations where individuals’ tendencies to cooperate are different, we have reported a new “level-off” phenomenon [28]; to be more specific, level-off here refers to the phenomenon that starting with a population of a low cooperation level, the level rises with the share of more-cooperative individuals but then levels off as the share reaches some threshold. These and other related simulation studies [29] provide intuition on how the decision-making dynamics in large populations governed by the best-response rule evolve over time, but sometimes the results are descriptive due to the lack of rigorous mathematical proofs. Some related mathematical results can in general only be applied to homogeneous populations in specific setups, such as deterministic updates within well-mixed and finite populations [30], infinite population or mean field approximation [31]–[34], and noisy updates [35]–[38]. Only a few papers with mathematical analysis have studied heterogeneous populations [39], [40], but still with restrictive assumptions like infinite populations or deterministic updates. So a great challenge is to develop new mathematically rigorous tools to analyze asynchronous decision-making dynamics under the best-response update rule in finite heterogeneous populations.

In this paper we consider a finite, well-mixed, heterogeneous population in which each individual, also referred to as agent, has her own (possibly unique) utility function when playing evolutionary games where the base game is the classical prisoner’s dilemma or snowdrift game to capture the well-known cooperation social dilemma. The agents are activated asynchronously to revise their strategies according to the best-response update rule. Our main results are to explicitly determine the long run dynamic behavior without using mean-field or other approximation methods. We are able to show that no matter what initial strategies the individuals decide to use, surprisingly one can always identify one type of individuals as a benchmark such that after a sufficiently long but finite time, more-cooperative individuals cooperate while those less cooperative defect. It is also shown that the total number of cooperators in the population will either become fixed or fluctuate at most by one. The importance of the new findings is threefold. First, these global convergence results explain rigorously why sustainable cooperation can emerge among selfish heterogeneous individuals. Second, the level-off phenomenon [28] is validated and proved rigorously and
thus shows the subtlety of controlling the level of cooperation if one wants to use population heterogeneity as a means to intervene collective decision-making processes. And third, the different thresholds in the level-off behavior for populations with diverse compositions indicate the sometimes restrictive nature of the homogeneity assumption in social evolutionary models.

The rest of the paper is organized as follows. In Section II, we present the model for decision-making dynamics. The main convergence results are shown in Section III. We in Section IV focus on the mathematical characteristics of the level-off phenomenon.

II. DECISION-MAKING UNDER BEST-RESPONSE UPDATES

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 activated to update her strategy according to how well she is doing when she plays her current strategy against the population. More specifically, the four possible payoffs of an agent \( i \), \( i = 1, \ldots, n \), are summarized in the \( 2 \times 2 \) payoff matrix

\[
A_i = \begin{pmatrix} C & D \\ D & T_i \end{pmatrix}.
\]

(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 agent \( i \)’s strategy at time \( t \) in the vector form, which is either \( s^C \triangleq [1 \, 0]^\top \) to cooperate or \( s^D \triangleq [0 \, 1]^\top \) to defect. Obviously, \( s^C = 1 - s^D \) with \( 1 = [1 \, 1]^\top \). So agent \( i \)’s payoff at time \( t \) against the population can be calculated by

\[
u_i(s_i(t), s_C(t)) = s_i(t)^\top A_i s_C(t)
\]

where \( s_C(t) \triangleq [n_C(t) - n_{-C}(t)]^\top \) and \( n_C(t) \) denotes the number of cooperators in the whole population at time \( t \). The myopic best-response update rule for agent \( i \) dictates that agent \( i \) chooses the strategy that maximizes her payoff. In case both cooperation and defection return the same payoff, we assume agent \( i \) sticks to her current strategy. Therefore, the update rule is that agent \( i \) sticks to her current strategy only if her alternative strategy does not give her a higher payoff, namely

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

(2)

As an illustration, the agents can be thought of as robots participating in a group transportation task, who can decide either to contribute to transport an object (C) or to free ride to let the other robots to do the job (D). Payoffs then represent the reward for accomplishing the task, minus the robot’s share for transformation cost. In a different scenario, the agents can also stand for drivers choosing to detour in order to reduce the potential traffic jam (C) or choosing the shortest path knowing this may increase the pressure for traffic management (D). Penalties can roughly be thought of as the time spent by each driver on the path, plus how much the driver is penalized to increase the potential traffic jam.

When studying homogeneous populations, people are mainly interested in how the portion of cooperators changes over time; for heterogeneous populations like in this paper, the more complex central topic is how the portions of cooperators with different utility functions evolve. Towards this end, we classify the heterogeneous individuals into different types according to their cooperating tendencies stipulated by their payoff matrices \( A_i \) in (1). We focus on those \( A_i \) with special structures [28]; the entries of \( A_i \) satisfy

\[
T_i > R_i > \max\{S_i, P_i\}, \quad S_i \neq P_i.
\]

(3)

Then \( A_i \) corresponds to either a prisoner’s dilemma (PD) game satisfying \( P_i > S_i \) or a snowdrift (SD) game satisfying \( S_i > P_i \). We call each agent with a PD (resp. SD) payoff matrix a PD agent (resp. SD agent). In fact, as we will later show in Lemma 1, PD agents always tend to defect under (2); however, an SD agent’s tendency to cooperate depends on the ratio \( \frac{S_i - P_i}{T_i - R_i + S_i - P_i} \), namely, the higher the ratio is, the more cooperative the agent becomes, except for those ratios in the same interval \( (\frac{k}{n}, \frac{k+1}{n}) \) for some \( k \in \{0, 1, \ldots, n - 1\} \) that result in the same cooperative tendency. For this reason, we categorize all those SD players whose ratios are within the same interval \( (\frac{k}{n}, \frac{k+1}{n}) \) into the same type. Moreover, all those SD players whose ratios equal \( \frac{k}{n} \), naturally fall into the same type as well. We assume there are altogether \( l > 0 \) types and label them by \( 1, \ldots, l \) according to the descending order of the ratios. Correspondingly, each SD agent of the \( j \)th, \( j \in \{1, \ldots, l\} \), type is called an \( S_{Dj} \) agent. Then there are altogether \( l + 1 \) types of agents after taking into account the PD agents. Let \( n_{PD} \) denote the number of PD agents and \( n_{SD} \), the number of \( S_{Dj} \) agents. Then the heterogeneity of the population is characterized by the type-vector

\[
p \triangleq (n_{SD1}, n_{SD2}, \ldots, n_{SDl}, n_{PD})
\]

and obviously

\[
p \in \mathbb{P}_n \triangleq \left\{ p \in \mathbb{Z}^{l+1} \left| \sum_{i=1}^{l+1} p_i = n, p_i \geq 0 \right. \right\}.
\]

We find that after scaling up the agents’ ratios by \( n \), some comparisons involving the numbers of different types of agents can be easier to be made. For this reason and to have the same number after scaling up the ratios of same-type agents, we define the temper of an \( S_{Dj} \) agent \( i \) to be

\[
n_{SDj}^* = \frac{1}{n} \left[ \frac{n}{T_i - R_i + S_i - P_i} + 1 \right] \frac{S_i - P_i}{T_i - R_i + S_i - P_i}
\]

(4)

where for a real number \( a \), \( [a] \) and \( \lceil a \rceil \) return the smallest integer no less than \( a \) and the largest integer no greater than \( a \) respectively. The temper of a PD agent is defined to be zero. We study the effects of heterogeneity of the population when similar types are sufficiently apart in tempers; to be more precise, we assume

\[
n_{SDj}^* > n_{SDj}^* \quad \text{and} \quad n_{SDj}^* > n_{SD(j+1)}^*, \quad \forall i \in \{2, \ldots, l-1\}.
\]

(5)
We use \( n_{SD}^C \) to denote the number of cooperators among the SD players and \( n_{PD}^C \) that among the PD players. Then we stack all the \( n_{SD}^C \) and \( n_{PD}^C \) together, and call the resulted \((l + 1)\)-dimensional vector the distribution of cooperators of the whole population

\[
x \triangleq (n_{SD1}^C, n_{SD2}^C, \ldots, n_{SDl}^C, n_{PD}^C).
\]

The main goal of this paper is to study given the type of a heterogeneous population \( p \in \mathcal{P} \), how the system state \( x \) evolves over time \( t \) under the update rule (2) and the activation sequence of the agents. Here, we assume that the agents are activated persistently as follows.

**Persistent activation assumption:** For every agent \( i \in \{1, \ldots, n\} \) and any time \( t = 0, 1, 2, \ldots \), there exists some finite time \( t' \geq t \) at which agent \( i \) is activated. This is a very mild assumption that holds almost surely, that is with probability one, in most of the usual stochastic settings, e.g., when each agent is activated at a rate determined by a Poisson clock [41].

Given the asynchronous, nonlinear nature of the dynamics of \( x(t) \), it is not clear whether \( x(t) \) will converge at all. In fact, we will show that depending on the given type \( p \), the state \( x(t) \) converges to either a single state or a set of two states, and more importantly the convergence takes place globally in finite time. The key technical step is to construct a Lyapunov-like function. We provide the details of our main convergence results and their proofs in the next section.

### III. Convergence of the Best-Response Dynamics

To understand better the best response update rule (2), we first rewrite it in the form emphasizing the effects of the agents’ tempers.

**Lemma 1:** When an agent \( i \in \{1, \ldots, n\} \) updates its strategy \( s_i(t) \) at time \( t \geq 0 \) according to (2), then

\[
s_i(t + 1) = s^D
\]

if \( i \) is a PD agent, and

\[
s_i(t + 1) = \begin{cases} 
  s^C & n_C(t) < n_{SDj}^*, \\
  s_i(t) & n_C(t) = n_{SDj}^*, \\
  s^D & n_C(t) > n_{SDj}^*,
\end{cases}
\]

if \( i \) is an SD\( j \) agent for some \( j \in \{1, \ldots, l\} \).

**Proof:** See subsection A in the Appendix.

Lemma 1 tells us that a PD agent will always update to defect while an SD agent’s decision depends on the comparison between its temper and the current total number of cooperators in the whole population. So the largest possible number of cooperators after every agent has updated at least once is \( n - n_{PD} = \sum_{j=1}^{l} n_{SDj} \). In this case, an immediate observation is that if

\[
n - n_{PD} < n_{SDi}^*,
\]

and even when all the possible cooperators (namely all the SD players) are indeed cooperating, the least cooperative type of SD player, i.e., SD\( i \) players, will all cooperate. This motivates us to distinguish populations with various \( p \)’s according to whether the inequality (8) holds.

**Definition 1:** For any of those \( p \) with which (8) holds, we say the corresponding population is **biased** since in it even the least cooperative type of SD agents tends to cooperate; otherwise, we say the population is **unbiased**.

We first observe that, from Lemma 1, if any PD agent is activated, her updated strategy is always to defect. Because of the persistent activation assumption, after a sufficient number of updates, all the PD agents have been activated at least once, then all of them stick to defection afterwards. We summarize it below.

**Lemma 2:** There exists a finite time \( t_{PD} \) such that

\[
x(t) = (n_{SD1}(t), n_{SD2}(t), \ldots, n_{SDl}(t), 0) \quad \forall t \geq t_{PD}.
\]

In the rest of the paper, we use \( t_{PD} \) in Lemma 2 to denote the time that the convergence to the state stipulated in the lemma has taken place. The long-run behavior for biased populations is rather straightforward to establish, which is presented in the following theorem.

**Theorem 1:** For a given biased population, there exists some time \( \tau \) such that for all \( t \geq \tau \)

\[
x(t) = (n_{SD1}, n_{SD2}, \ldots, n_{SDl}, 0).
\]

**Proof:** For a biased population, it holds that

\[
n_C(t) \leq \sum_{j=1}^{l} n_{SDj} \quad \forall t \geq t_{PD}
\]

(8): \( \sum_{j=1}^{l} n_{SDj} = n - n_{PD} < n_{SD}^* \)

(5): \( n_{SDi}^* \leq n_{SDi} \quad \forall i \in \{1, \ldots, l\} \)

\[
n_C(t) < n_{SDi}^* \quad \forall i \in \{1, \ldots, l\}, \forall t \geq t_{PD}.
\]

Hence, from Lemma 1, if any of the SD agents is activated after \( t_{PD} \), she chooses cooperation as her strategy. Furthermore, since the agents are activated persistently, there exists some time \( \tau \geq t_{PD} \) such that all of the SD agents have become activated before \( \tau \), and thus choose cooperation without changing it afterwards. So the proof is complete.

However, the analysis of the long-run behavior for unbiased populations is much more complicated, and we need some additional technical notions. For an unbiased population, we can still examine the accumulated number of SD players according to their ordering, and more precisely we define the benchmark

\[
k_p \triangleq \min \left\{ k \mid k \in \{1, \ldots, l\}, \sum_{i=1}^{k} n_{SDi} \geq n_{SDk}^* \right\}
\]

and call the SD\( k_p \) agents the **benchmark agents** since such players, as will be shown later, indicate clearly who will always cooperate in the long run. The following property follows directly from this definition of \( k_p \):

\[
\sum_{i=1}^{j} n_{SDi} < n_{SDj}^* \quad \forall j < k_p.
\]

Now we are ready to present the first main result of this section, which identifies the invariant set of the system for
unbiased populations. To do so, we define the following \((l+1)\)-dimensional vectors
\[
x^* \triangleq (n_{SD_1}, \ldots, n_{SD_{kp-1}}, 0, \ldots, 0),
\]
\[
x^- \triangleq (n_{SD_1}, \ldots, n_{SD_{kp-1}}, [n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}, 0, \ldots, 0),
\]
\[
x^+ \triangleq (n_{SD_1}, \ldots, n_{SD_{kp-1}}, [n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}, 0, \ldots, 0).
\]

Note that these three vectors differ only in their \(kp\)th element; such elements in \(x^+\) and \(x^-\) in general differ by one, and are the same if and only if \([n^*_{SD_{kp}}]\) is an integer. For unbiased populations, we further classify them into two categories: we say the population is clean-cut if \(\sum_{j=1}^{kp-1} n_{SD_j} \geq [n^*_{SD_{kp}}]\); otherwise, ruffled.

**Proposition 1:** For a given unbiased population with the benchmark \(kp\), when the population is clean-cut, \(x^*\) is invariant; otherwise when it is ruffled, \([x^-, x^+]\) is.

We need the following lemma in the proof.

**Lemma 3:** If \(x(t_0) \in \{x^*, x^+, x^-\}\) at some \(t_0 \geq t_{PD}\), then
\[
n^{C}_{SD_j}(t_0 + 1) = n^{C}_{SD_j}(t_0) = 0 \quad \forall i > kp, \quad (11)
\]
\[
n^{C}_{SD_j}(t_0 + 1) = n^{C}_{SD_j}(t_0) = n_{SD_i} \quad \forall i < kp. \quad (12)
\]

**Proof:** See Subsection B in the Appendix. 

**Proof of Proposition 1:** It suffices to show that for some \(t_0 \geq t_{PD}\), in clean-cut populations
\[
(x(t_0) = x^*) \Rightarrow (x(t) = x^* \quad \forall t \geq t_0), \quad (13)
\]
and in ruffled populations
\[
(x(t_0) \in \{x^-, x^+\}) \Rightarrow (x(t) \in \{x^-, x^+\} \quad \forall t \geq t_0). \quad (14)
\]
We first observe that in a clean-cut population at \(x^*\),
\[
n^{C}_{SD_{kp}}(t_0) = 0. \quad \text{Then from (46) and Lemma 1, we know that}
\]
\[
n^{C}_{SD_{kp}}(t_0 + 1) = n^{C}_{SD_{kp}}(t_0) = 0. \quad (15)
\]
Combining (15) with (11) and (12) in Lemma 3, we conclude that (13) holds for \(t = t_0 + 1\) and then by induction holds for any \(t \geq t_0\). We then observe that in a ruffled population at \(x^-\) and \(x^+\),
\[
n^{C}_{SD_{kp}}(t_0) \in \{[n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}, [n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}\}. \quad (16)
\]
Then one of the following three cases must take place.

**Case 1:** \([n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j} = [n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}\). Then \([n^*_{SD_{kp}}] = [n^*_{SD_{kp}}] = n^*_{SD_{kp}}\), implying that \(n^C(t_0) = n^*_{SD_{kp}}\). Hence, from Lemma 1, if an \(SD_{kp}\) agent is active at \(t_0\), she will not change her strategy. Therefore,
\[
n^{C}_{SD_{kp}}(t_0 + 1) = n^{C}_{SD_{kp}}(t_0) \quad (17)
\]

**Case 2:** \([n^*_{SD_{kp}}] = [n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}\). Then
\[
n^C(t_0) = n^*_{SD_{kp}} > n^*_{SD_{kp}}.
\]
Hence, from Lemma 1, if an \(SD_{kp}\) agent is active at \(t_0\), she updates to \(s^C\) at \(t_0 + 1\). So
\[
n^{C}_{SD_{kp}}(t_0 + 1) = n^{C}_{SD_{kp}}(t_0) - 1 = [n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}. \quad (18)
\]

**Case 3:** \([n^*_{SD_{kp}}] = [n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}\). Then
\[
n^C(t_0) = [n^*_{SD_{kp}}] < n^*_{SD_{kp}}.
\]
Hence, from 1, when an \(SD_{kp}\) agent is active at \(t_0\), she updates to \(s^C\) at \(t_0 + 1\). So
\[
n^{C}_{SD_{kp}}(t_0 + 1) = n^{C}_{SD_{kp}}(t_0) + 1 = [n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}. \quad (19)
\]

Moreover, when the active agent at \(t_0\) is not an \(SD_{kp}\) agent, we have that
\[
n^{C}_{SD_{kp}}(t_0 + 1) = n^{C}_{SD_{kp}}(t_0). \quad (20)
\]
Combining (17), (18), (19) and (20) together, we have
\[
n^{C}_{SD_{kp}}(t_0+1) \in \{[n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}, [n^*_{SD_{kp}}] - \sum_{j=1}^{kp-1} n_{SD_j}\}. \quad (21)
\]
Combining the above deduction with (11) and (12) in Lemma 3, we conclude that (14) holds for \(t = t_0 + 1\) and then by induction holds for any \(t \geq t_0\).

In fact, a much stronger statement than Proposition 1 can be made, which shows the invariant state/set is also globally attractive.

**Theorem 2:** For a given unbiased population with the benchmark \(kp\), if the population is clean-cut, then there exists some time \(\tau\) such that
\[
x(t) = x^* \quad \forall t \geq \tau; \quad (22)
\]
otherwise, there exists some time \(\tau\) and two infinite time sequences \(\{t^-\}_{i=1}^{\infty}\) and \(\{t^+\}_{i=1}^{\infty}\) such that
\[
x(t) \in \{x^-, x^+\} \quad \forall t \geq \tau, \quad (22)
\]
and
\[
x(t^-) = x^-, \quad x(t^+) = x^+ \quad i = 1, 2, \ldots.
\]

The theorem implies that after finite time the population reaches a state where every \(SD_i\), \(i = 1, \ldots, kp-1\), agent cooperates, and every \(SD_i\), \(i = kp + 1, \ldots, l\), agent and every \(PD\) agent defects. In other words, all the agents that have a higher (resp. lower) temper than that of the benchmark agents, become cooperators (resp. defectors) after a sufficiently long, but finite, time. Moreover, if the population is clean-cut, all \(SD_{kp}\) agents defect. Otherwise, if further \(x^* = x^+\) or equivalently \([n^*_{SD_{kp}}] = [n^*_{SD_{kp}}]\), we have that
\[
n^*_{SD_{kp}} = \sum_{j=1}^{kp-1} n_{SD_j} \quad \text{of the \(SD_{kp}\) agents cooperate, and the}
rest defect; else in case \( x^* \neq x^*_+ \), the number of \( SD_kp \) agents who cooperate fluctuates between \( |n^*_SDk_p| - \sum_{j=1}^{k_p-1} nSD_j \) and \( |n^*_SD_{k_p}| - \sum_{j=1}^{k_p-1} nSD_j \).

To prove Theorem 2, we make use of a Lyapunov-like function \( h(x) \) defined for populations with \( k_p \geq 2 \). Let \( \alpha(x) = 0 \) if \( n^C_{SD_1} \neq nSD_1 \); otherwise, let \( \alpha(x) \) return the largest index of those \( SD \) agents such that

\[
\begin{align*}
n^C_{SD_j}(t) &= nSD_j \quad \forall j \in \{1, \ldots, \alpha(x)\}, \\
\alpha(x) &\leq k_p - 2.
\end{align*}
\]

Then \( h(x) \) is defined to be

\[
h(x) \triangleq \sum_{i=1}^{\alpha(x)+1} n^C_{SD_i}.
\]

Obviously, \( h(x) \) is lower bounded by zero and upper bounded by \( \sum_{j=1}^{k_p-1} nSD_j \). One main step in proving Theorem 2 is to show that \( h(x) \)'s upper bound is tight and along the system's trajectory \( x(t) \), \( h \) reaches and stays at its maximum after finite time, after which all \( SD_1, SD_2, \ldots, SD_{k_p-1} \) agents always cooperate. Then by using a similar function \( g(x) \), we can prove all \( SD_{k_p+1}, SD_{k_p+2}, \ldots, SD_l \) agents always defect after finite time, and the strategies of the \( SD_k \) players depends on whether the inequality \( \sum_{j=1}^{k_p-1} nSD_j \geq nSD_{k_p} \) holds. Although along \( x(t) \), \( h \) does not always increase, its lower bound can be tightened up as time grows. To be more specific, we will construct an infinite set \( B \) of time instants \( t_b \), such that for any \( t_b \in B \), it holds that \( h(t) \geq h(t_b) \) for all \( t \geq t_b \); more precisely

\[
B = \{ t_b \} \triangleq \left\{ t \geq t_{PD}, \quad nC(t) = \sum_{j=1}^{k_p-1} nSD_j \text{ if clean-cut} \right. \\
\text{or } nC(t) \in \{[n^*_SD_{k_p}], [n^*_SD_{k_p}] \} \text{ otherwise} \}.
\]

Then at \( t_b \), the number of cooperators in a clean-cut population is \( x^* \), and that in a ruffled population is \( x^*_+ \) or \( x^*_+ \). Now we show by the following two lemmas that such a constructed \( B \) is not only non-empty but infinite.

**Lemma 4:** In a clean-cut population, if at some time \( T > t_{PD} \),

\[
nC(T) < \sum_{j=1}^{k_p-1} nSD_j,
\]

then there exists a finite time \( T' > T \) such that

\[
nC(T') = \sum_{j=1}^{k_p-1} nSD_j.
\]

In a ruffled population, if at some time \( T > t_{PD} \),

\[
nC(T) < [n^*_SD_{k_p}],
\]

then there exists some finite time \( T' \) such that

\[
nC(T') \in \{[n^*_SD_{k_p}], [n^*_SD_{k_p}] \}.
\]

**Proof:** For the clean-cut population, we first prove by contradiction that there exists some finite time \( T > T \) such that

\[
nC(T) \geq \sum_{j=1}^{k_p-1} nSD_j.
\]

Assume the contrary, that is

\[
nC(t) < \sum_{j=1}^{k_p-1} nSD_j, \quad \forall t > T.
\]

Then from (10), for all \( t > T \) we have that

\[
nC(t) < n^*_SD_{k_p} \quad \forall i < k_p.
\]

Hence, from Lemma 1 and the persistent activation assumption, we know that there exists a finite time \( \tau \), before which all of the \( SD_i, i < k_p \), agents, have updated at least once to fix their strategies to cooperation. Then

\[
x(\tau) = (nSD_1, \ldots, nSD_{k_p-1}, n^C_{SD_{k_p}}(\tau), \ldots, n^C_{SD_1}(\tau), 0),
\]

where the zero component follows from Lemma 2, implying

\[
nC(\tau) = \sum_{j=1}^{k_p-1} nSD_j(\tau) \geq \sum_{j=1}^{k_p-1} nSD_j,
\]

which contradicts (30). So (29) must be true.

Further, since \( nC \) takes integer values and changes at most by one each time, the transition of \( nC \) from (25) to (29) implies the existence of \( T' \) to make \( nC \) become \( \sum_{j=1}^{k_p-1} nSD_j \) at \( T' \), and thus (26) is proved.

For the ruffled population, similarly, we first prove by contradiction the existence of a finite time \( T > T \) such that

\[
nC(T) \geq [n^*_SD_{k_p}].
\]

Assume the contrary, that is

\[
nC(t) < [n^*_SD_{k_p}] \quad \forall t > T.
\]

From \( [n^*_SD_{k_p}] \leq n^*_SD_{k_p} \), it follows that for all \( t > T \),

\[
nC(t) < n^*_SD_{k_p} \quad \forall i \leq k_p.
\]

Then from Lemma 1, the persistent activation assumption and Lemma 2, there is a finite time \( \tau > T \) such that

\[
x(\tau) = (nSD_1, \ldots, nSD_{k_p}, n^C_{SD_{k_p-1}}(\tau), \ldots, n^C_{SD_1}(\tau), 0).
\]

Then because of the definition of \( k_p \),

\[
nC(\tau) = \sum_{j=1}^{k_p} nSD_j(\tau) \geq \sum_{j=1}^{k_p} nSD_j \geq n^*_SD_{k_p} \geq [n^*_SD_{k_p}],
\]

which contradicts (32). So (31) is true and thus the transition of \( nC \) from (27) to (31) implies (28).

**Lemma 5:** In a clean-cut population, if at \( T > t_{PD} \),

\[
nC(T) > \sum_{j=1}^{k_p-1} nSD_j.
\]

In a ruffled population, if at some time \( T > t_{PD} \),

\[
nC(T) < [n^*_SD_{k_p}],
\]

then there exists some finite time \( T' \) such that

\[
nC(T') \in \{[n^*_SD_{k_p}], [n^*_SD_{k_p}] \}.
\]
then there exists a finite time $T' > T$ such that
\[
    n_C(T') = \sum_{j=1}^{k_p-1} n_{SD_j}.
\]
(34)

In a ruffled population, if at $T > t_{PD}$,
\[
    n_C(t) > \left[ n_{SD_{kp}}^* \right],
\]
(35)
then there exists a finite time $T'$ such that
\[
    n_C(T') \in \left\{ \left[ n_{SD_{kp}}^* \right], \left[ n_{SD_{kp}}^* \right] \right\}.
\]
(36)

This lemma can be proven by contradiction following the analogous statements in the proof for Lemma 4. Since Lemma 4 covers the situation for $T > t_{PD}$, $n_C(t) < \sum_{j=1}^{k_p-1} n_{SD_j}$, Lemma 5 for the situation $n_C(t)$ greater than the sum, and the situation $n_C$ equals the sum corresponds to $t \in B$, we have shown that for any $T > t_{PD}$, either $t \in B$ or there is a finite $T' > T$ satisfying $T' \in B$. So we have actually proven the following.

**Lemma 6:** The set $B$ is infinite.

After knowing $B$ has infinite time instants in it, we show the nondecreasing property of $h(x)$ along $x(t)$ at $t \in B$.

**Lemma 7:** If $t_b \in B$, then
\[
    h(t) \geq h(t_b) \quad \forall t \geq t_b.
\]
(37)

We need the following two lemmas in the proof.

**Lemma 8:** For all $t_b \in B$, it holds that
\[
    n_C(t_b) < n_{SD_{kp-1}}^*.
\]
(38)

**Proof:** See Subsection C in the Appendix.

**Lemma 9:** If $t_b \in B$, then
\[
    n_C(t) < n_{SD_{\alpha(t)+1}}^* \quad \forall t \geq t_b.
\]
(39)

**Proof:** See Subsection D in the Appendix.

**Proof of Lemma 7:** We prove by contradiction. Assume on the contrary that (37) can be violated at some time $t_1 \geq t_b+1$. Hence, for some $i = 1, \ldots, n_{SD_{\alpha(t)+1}}$, an $SD_i$ agent has changed its choice from $C$ to $D$ at $t_1$. So from Lemma 1,
\[
    n_C(t_1-1) > n_{SD_i}.
\]
(40)

On the other hand, it follows from (5) that $n_{SD_i} \geq n_{SD_{\alpha(t)+1}}^*$. Hence, (40) implies
\[
    n_C(t_1-1) > n_{SD_{\alpha(t)+1}}^*,
\]
which contradicts Lemma 9 since $t_1 - 1 \geq t_b$.

**Lemma 10:** For a fixed $t_b \in B$, consider the first time after $t_b$ that an $SD_{\alpha(t)+1}$ agent, whose strategy was $s^D$ at $t_b$, is activated. Then this agent updates her strategy to $s^C$, and does not change it afterwards.

Then we prove that $h$ reaches its maximum.

**Lemma 11:** There exists some time $t_h$, at which $h$ reaches and after which remains at its maximum, i.e., $h(t) = \sum_{j=1}^{k_p-1} n_{SD_j}$ for all $t \geq t_h$.

**Proof:** If there exists $t_b \in B$ such that $h(t_b) = \sum_{j=1}^{k_p-1} n_{SD_j}$, then from Lemma 7, $h(t) \geq \sum_{j=1}^{k_p-1} n_{SD_j}$ for all $t \geq t_b$, and thus the result holds straightforwardly. Otherwise, the only other possibility is that for each $t_b \in B$, $h(t_b) < \sum_{j=1}^{k_p-1} n_{SD_j}$. We will show below that this, however, will never happen. In this case, for every $t_b \in B$, there exists an $SD_{\alpha(t)+1}$ agent whose strategy is $s^D$ at $t_b$. On the other hand, due to the persistent activation assumption, every such agent will become activated at some finite future time, and in view of Lemma 11, will update her choice to $C$ and will not change it afterwards. This cannot be true since $B$ is finite according to Lemma 6, but the number of $SD_{\alpha(t)+1}$ agents are finite. So the contradiction completes the proof.

**Lemma 12:** There exists some time $t_g$ at which $g$ reaches and remains at its minimum, i.e., $g(t) = 0$ for all $t \geq t_g$.

We are ready to prove the main result.

**Proof of Theorem 2:** If $k_p \geq 2$, then in view of Lemma 11 and according to the definition of $h$, there exists some time $t_h$ such that
\[
    n_{SD_j}(t) = n_{SD_j} \quad \forall t \geq t_h, \forall j \in \{1, \ldots, k_p-1\}.
\]
(41)

On the other hand, if $k_p \leq l - 1$, then in view of Lemma 12 and according to the definition of $g$, there exists some time $t_g$ such that
\[
    n_{SD_j}(t) = 0 \quad \forall t \geq t_g, \forall j \in \{k_p+1, \ldots, l\}.
\]
(42)

Let $t_{max} = \max \{ t_h, t_g \}$. From (41) and (42), it follows
\[
    x(t) = (n_{SD_1}, \ldots, n_{SD_{k_p-1}}, n_{SD_{k_p}}(t), 0, \ldots, 0) \quad \forall t \geq t_{max}.
\]
(43)

Clearly (43) also holds if $k_p = 1$ or $k_p = l$, implying that it holds in general for any unbiased population. Now consider $t_b \in B$, $t_b \geq t_{max}$. If the population is clean-cut, then by the definition of $B$,
\[
    n_C(t_b) = \sum_{j=1}^{k_p-1} n_{SD_j} \Rightarrow n_{SD_{kp}}^*(t_b) = n_C(t_b) - \sum_{j=1}^{k_p-1} n_{SD_j} = 0.
\]

Hence, $x(t_b) = x^*$, and in view of Lemma 1, $x(t) = x^*$ for all $t \geq t_b$. On the other hand, if the population is ruffled, then
\[
    n_C(t_b) \in \left\{ \left[ n_{SD_{kp}}^* \right], \left[ n_{SD_{kp}}^* \right] \right\} \Rightarrow
\]
(43)
\[
    n_{SD_{kp}}^*(t_b) = n_C(t_b) - \sum_{j=1}^{k_p-1} n_{SD_j} \in \left\{ \left[ n_{SD_{kp}}^* \right] - \sum_{j=1}^{k_p-1} n_{SD_j}, \left[ n_{SD_{kp}}^* \right] - \sum_{j=1}^{k_p-1} n_{SD_j} \right\}.
\]
Hence, \( x(t_b) \in \{ x^\ast, x^* \} \), and in view of Lemma 1, \( x(t) \in \{ x^\ast, x^* \} \) for all \( t \geq t_b \). Therefore, (21) and (22) are proven by taking \( \tau = t_b \).

To construct the two infinite subsequences, if \( x^* = x^* \), the result is trivial. Otherwise, let \( \tau_i > \tau \) be the ith time after \( t = \tau \) that an \( SD_k \) agent is activated. The time sequence \( \{ \tau_i \} \) is infinite since the agents are activated persistently. From (22), \( x(\tau_i) \in \{ x^\ast, x^* \} \). Without loss of generality assume \( x(\tau_i) = x^* \). From the definition of \( x^* \),

\[
n_C(\tau_i) = \sum_{j=1}^{k_{\ast}-1} n_{SD_i} + n_{SD_k}(\tau_i) = [n_{SD_k}]_n. \tag{44}
\]

It can be easily verified that \( x(t) \) does not change within \([\tau_i, \tau_{i+1}]\). Hence, \( n_C(\tau_{i+1}) = n_C(\tau_i) \). On the other hand, \( x^* \neq x^* \) yields \( n_{SD_k}(\tau_i) < n_{SD_k} < [n_{SD_k}]_n \). Hence, because of (44), \( n_{SD_k}(\tau_i) < n_{SD_k} \), which is in view of Lemma 1 implies that the active \( SD_k \) agent at \( \tau_i \) switches her choice from \( D \) to \( C \). Hence, \( n_{SD_k}\{ \tau_i \} = n_{SD_k}(\tau_i) + 1 = [n_{SD_k}]_n - \sum_{j=1}^{k_{\ast}-1} n_{SD_i}. \)

Therefore, \( x(\tau_{i+1}) = x^* \). Similarly, it can be shown that \( x(\tau_{i+2}) = x^* \) for any integer. Then

\[
x(\tau_{2i+1}) = x^*, \quad x(\tau_{2i+2}) = x^* \quad \forall r = 0, 1, 2, \ldots.
\]

Taking \( \{ t_i \}_{i=1}^{\infty} = \{ \tau_{i+1} \}_{r=0}^{\infty} \) and \( \{ t_i \}_{i=1}^{\infty} = \{ \tau_{2i} \}_{r=0}^{\infty} \), we arrive at the conclusion.

Remark 1: Theorem 2 also holds when condition (5) is relaxed to the following

\[ n_{SD_i}^\ast > n_{SD_i}^\ast, \quad \forall i \in \{1, \ldots, l-1\}; \]

however, then the agents must be activated pairwise persistently, that is for any pair of agents \( i \) and \( j \) and each time \( t \), there exists some finite time \( t' > t \) such that \( i \) and \( j \) are activated consecutively at \( t' \) and \( t' + 1 \), respectively. Although stronger than the persistent activation assumption, the pairwise persistent activation assumption is satisfied almost surely in most stochastic settings as well, particularly when agents are activated independently, e.g., according to Poisson clocks.

The following result can be derived directly from Theorems 1 and 2, and specifies the number of long-run cooperators.

Corollary 1: For a population with \( p \in \mathbb{P}_n \), there exists some time \( \tau \) such that for all \( t \geq \tau \), if the population is biased, \( n_C(t) = \sum_{i=1}^{n} n_{SD_i} \); otherwise, if it is clean-cut,

\[ n_{SD_i} - n_{SD_i}^\ast, \quad \forall i \in \{1, \ldots, l-1\}; \]

and if it is ruffled,

\[ n_{SD_i}, \quad \forall i \in \{1, \ldots, l-1\}; \]

and if it is ruffled,

\[ n_{SD_i} - n_{SD_i}^\ast, \quad \forall i \in \{1, \ldots, l-1\}; \]

and if it is ruffled,

\[ n_{SD_i}^\ast, \quad \forall i = 1, 2, \ldots, \]

\[ n_{SD_i} = [n_{SD_i}^\ast], \quad n_{SD_i}^\ast = [n_{SD_i}^\ast]. \]

Corollary 1 and Lemma 8 can be used to derive lower and upper bounds for the number of long-run cooperators.

Corollary 2: For an unbiased population with benchmark \( k_p \), there exists some time \( \tau \) such that for all \( t \geq \tau \),

\[ n_{SD_k} \leq n_{SD_i} - n_{SD_i}^\ast, \quad \forall i \in \{1, \ldots, l-1\}. \]

After having presented the mathematical analysis of the long-run behavior of the decision-making dynamics, in the next section, we look into how the number of long-run cooperators changes when the cooperation tendencies of some \( SD \) players are manipulated.

IV. THE LEVEL-OFF PHENOMENON

For heterogeneous populations, important research questions arise that do not show up for homogeneous populations. In particular, we are interested in knowing whether increasing some individuals’ tendencies to cooperate results in the rise of the cooperation level of the whole population. Mainly through simulations, we have provided a negative answer to this question in [28]. In this paper, we give more thorough theoretical analysis showing that the total number of cooperators in the long run, denoted by \( n_{SD}^f \), may initially increase as the tendencies of a portion of individuals increases, but levels off as the portion grows until the size of the portion passes a threshold, after which the number of cooperators may continue to grow. We call this particular relationship between the number of long-run cooperators and the portion of manipulated population the level-off phenomenon. Let us first illustrate this phenomenon by an example.

Example 1: Consider the decision-making dynamics formulated in Sec. II for a population of 100 agents with six different types, namely \( n = 100 \) and \( l = 5 \). Set the temps to be

\[ (n_{SD_1}, n_{SD_2}, n_{SD_3}, n_{SD_4}, n_{SD_5}) = (75, 60, 50, 40, 15), \]

and the type-vector

\[ p = (7, 0, 3, 20, 60, 10). \]

For \( r = 0, 1, \ldots, 60 \) agents of the \( SD_5 \) type, we change them to \( SD_3 \) agents. We plot \( n_{SD}^f \) against \( r \) in Figure 1. Clearly, as the number of \( SD_3 \) agents increases, the number of long-run cooperators \( n_{SD}^f \) first increases as \( r \) grows from 0 to 10, levels off for 10 \( \leq r \leq 30 \), increases again for 30 \( \leq r \leq 40 \), and experiences two more level-offs afterwards. So increasing \( r \) does not necessarily increase the number of long-run cooperators.

We remark that the temps in Example 1 are chosen to be integers to avoid the outcome that the total number of cooperators fluctuates between two consecutive numbers in the long run. This simplifies the plot in Figure 1 and the presentation of its corresponding mathematical results, which motivates us to focus on integer temps in what follows. When the temps are allowed to be real numbers, the total number of cooperators in the long run may fluctuate between two consecutive integers at some values of \( n_{SD} \) in Figure 1; however, the overall shape of the curve remains unchanged.

To explain such level-off phenomena mathematically, we first observe from Figure 1 that every plateau in the curve takes the value of one of the integer-valued temps \( n_{SD}^b, b \in \{1, \ldots, l\} \). We formulate this observation more rigorously in the following proposition.

**Proposition 2:** For \( b \in \{1, \ldots, l\} \), assume \( n_{SD}^b \) is an integer. Then \( n_{SD}^f = n_{SD}^b \), if and only if \( \sum_{s=1}^{b-1} p_s \leq n_{SD}^f \leq \sum_{s=1}^{b} p_s \).
Fig. 1. For Example 1, the number of long-run cooperators $n_C^f$ against the number of $SD_2$ agents $r$, whose types have changed to $SD_2$.

Proof: (sufficiency) Since

$$\sum_{s=1}^{b} p_s \leq n_{SD_b}^* < n_{SD_{b-1}}^* \quad \text{(5)}$$

$$\sum_{s=1}^{b} p_s \geq n_{SD_b}^*$$

it follows from Corollary 1 and $\sum_{s=1}^{b} p_s \leq n_{SD_b}^*$ that $n_C^f(p) = n_{SD_b}^*$. (necessity) Let $n_C^f(p) = n_{SD_b}^*$. First $n_{SD_b}^* \leq \sum_{s=1}^{b} p_s$, then $p_s \geq b+1$. Consider the situation when the population is clean-cut, i.e., $n_{SD_b}^* = \sum_{s=1}^{b} p_s$, then $k_p \geq b+1$. Hence, in view of Corollary 1 and the definition of $k_p$, $\sum_{s=1}^{b} p_s < n_{SD_{b-1}}^*$. Therefore, the population is ruffled, i.e., $\sum_{s=1}^{b} p_s < n_{SD_{b-1}}^*$. Hence, in view of Corollary 1, $n_C^f(p) = n_{SD_b}^* = n_{SD_{b-1}}^*$, this is in contradiction with $k_p \geq b+1$. So $n_{SD_b}^* \leq \sum_{s=1}^{b} p_s$ must be true.

Now $\sum_{s=1}^{b} p_s \leq n_{SD_b}^*$ is proven by contradiction. Assume the contrary, i.e., $\sum_{s=1}^{b} p_s > n_{SD_b}^*$. If $\sum_{s=1}^{b} p_s < n_{SD_{b-1}}^*$, then $p_s = b$, and in view of Corollary 1, $n_C^f(p) = \sum_{s=1}^{b} p_s > n_{SD_b}^*$, which is impossible. Hence, $\sum_{s=1}^{b} p_s \geq n_{SD_b}^*$, yielding $k_p \leq b-1$. On the other hand, in view of Corollary 2, $n_C^f(p) \geq n_{SD_{b-1}}^*$, which according to (5) and $k_p \leq b-1$ results in $n_C^f(p) > n_{SD_b}^*$, a contradiction. Hence, $\sum_{s=1}^{b} p_s \leq n_{SD_b}^*$.

The proposition shows the existence of a range of type-vectors that will lead to $n_{SD_b}^*$ cooperators in the population in the long run. The following theorem takes advantage of this result, and determines precisely those changes in the types of the agents in order to achieve and maintain $n_{SD_b}^*$ cooperators in the long run. For a type-vector $p$, denote by $p_{i\rightarrow j}$, $i, j \in \{1, \ldots , l\}$, the new type-vector after the type of $r \geq 0$ agents changes from $SD_i$ to $SD_j$. When all the tempers $n_{SD_b}^*$ are integers, from Corollary 1, we know that $n_C^f(p)$ always exists for any $p \in P_n$.

**Theorem 3 (sustainable cooperation levels):** For a given population with $p \in P_n$, assume all the tempers $n_{SD_b}^*$ are integers and assume there exists some $b \in \{1, \ldots , l\}$ such that $n_C^f(p) = n_{SD_b}^*$. Then

1. $n_C^f(p_{i\rightarrow j}) > n_C^f(p)$, if and only if $i \geq b, j \leq b-1$ and $r > n_{SD_b}^* - \sum_{s=1}^{b-1} p_s$.
2. $n_C^f(p_{i\rightarrow j}) < n_C^f(p)$, if and only if $i \leq b, j \geq b+1$ and $r < \sum_{s=1}^{b} p_s - n_{SD_b}^*$.
3. $n_C^f(p_{i\rightarrow j}) = n_C^f(p)$, otherwise.

Proof: Using Proposition 2, it can be shown that $n_C^f(p_{i\rightarrow j}) = n_C^f(p)$, if and only if one of the following holds

1. $i = j$.
2. $i \geq b+1$ and $j \geq b$.
3. $i \geq b+1, j \leq b-1$ and $r \leq n_{SD_b}^* - \sum_{s=1}^{b-1} p_s$.
4. $i = b, j \leq b-1$ and $r \leq n_{SD_b}^* - \sum_{s=1}^{b} p_s$.
5. $i = b, j > b+1$ and $r \leq \sum_{s=1}^{b} p_s - n_{SD_b}^*$.
6. $i \leq b-1$ and $j \leq b$.
7. $i \leq b-1, j \geq b+1$ and $r \leq \sum_{s=1}^{b} p_s - n_{SD_b}^*$.

Hence, in order to have $n_C^f(p_{i\rightarrow j})$ greater than (resp. less than) $n_C^f(p)$, all of the above cases must be violated. This results in either $i \geq b, j \leq b-1$ and $r \geq n_{SD_b}^* - \sum_{s=1}^{b-1} p_s$ or $i \leq b, j \geq b+1$ and $r \geq \sum_{s=1}^{b} p_s - n_{SD_b}^*$. It can be verified that only in the first case, $n_C^f(p_{i\rightarrow j})$ becomes greater than $n_C^f(p)$, and only in the second case, $n_C^f(p_{i\rightarrow j})$ becomes less than $n_C^f(p)$. Hence, in any other case, $n_C^f(p_{i\rightarrow j}) = n_C^f(p)$, which completes the proof.

Theorem 3 confirms the existence of cooperation levels that are robust against changes in the cooperation tendencies of the agents. Namely, unless one of the first two cases in the theorem takes place, the number of long-run cooperators is robust against the changes in types of the population. Note that this robustness is against both increasing and decreasing cooperation tendencies.

Another usage of Theorem 3 is to determine quantitatively the widths of the plateaus for curves like Figure 1. In Example 1, since $i \geq b$ and $j \leq b-1$, Case 1) of Theorem 3 implies that increasing the type of $r SD_b$ agents to $SD_2$, does not increase the total number of cooperators if and only if $r \leq n_{SD_b}^* - \sum_{s=1}^{3} p_s = 40 - 20 = 20$. Therefore, the width of the first plateau is 20.

The following proposition discusses what happens when the number of cooperators is not equal to any temper $n_{SD_b}^*$.

**Proposition 3:** Given a population with $p \in P_n$, if the number of long-run cooperators is some constant between two consecutive tempers of the agents, $n_{SD_b}^*$ and $n_{SD_{b-1}}^*$, $b \in \{2, \ldots , l\}$, where $n_{SD_{b-1}}^*$ is an integer, then changing an $SD_b$ agent to $SD_j$ where $j < b \leq i$, increases the number of long-run cooperators, i.e.,

$$n_{SD_b}^* < n_C^f(p) < n_{SD_{b-1}}^* \Rightarrow n_C^f(p_{i\rightarrow j}) > n_C^f(p), \; j < b \leq i.$$
Proof: From Corollary 1, the condition \( n_{SD_b}^* < n_{SD}^* (p) < n_{SD_{b-1}}^* \) implies \( n_C^f (p) = \sum_{s=1}^{b-1} p_s \geq n_{SD_{b-1}}^* \). Then, since \( \sum_{s=1}^{b-1} p_s < n_{SD_{b-1}}^* \), it holds that \( k_p = b \). Let \( \hat{p} = p_{i \rightarrow j}^f \). Clearly \( \sum_{s=1}^{b-1} \hat{p}_s = \sum_{s=1}^{b-1} p_s = \sum_{s=1}^{b-1} p_s + 1 \). Hence,

\[
n_{SD_b}^* < \sum_{s=1}^{b-1} \hat{p}_s \leq n_{SD_{b-1}}^*.
\]

If \( \sum_{s=1}^{b-1} \hat{p}_s < n_{SD_{b-1}}^* \), then \( \hat{p} = b \) and the population is clean-cut. So in view of Corollary 1, \( n_{SD_b}^* (\hat{p}) = \sum_{s=1}^{b-1} \hat{p}_s = \sum_{s=1}^{b-1} p_s + 1 = n_C^f (p) + 1 \). If on the other hand, \( \sum_{s=1}^{b-1} \hat{p}_s = n_{SD_{b-1}}^* \), then it can be shown that \( k_p = b - 1 \). Hence, from Corollary 2 and since \( n_{SD_{b-1}}^* \) is an integer, we have that \( n_{SD_b}^* (\hat{p}) = n_{SD_{b-1}}^* > n_{SD}^* (p) \).

As shown in the proof of Proposition 3, the condition \( n_{SD_b}^* < n_{SD}^* (p) < n_{SD_{b-1}}^* \) implies \( k_p = b \) and \( \sum_{s=1}^{b-1} p_s \geq n_{SD_{b-1}}^* \). Hence, from Theorem 2, the condition \( j < b \leq i \) in Proposition 3 implies that a defector changes to a cooperator. On the other hand, in view of Theorem 3, the above proposition implies that sustainable cooperation levels only maintain at the tempers of the agents. Therefore, Proposition 3 together with Theorem 3 provide a complete characterization of the level-off phenomena.

V. CONCLUDING REMARKS

We have studied a finite heterogeneous population of decision-making agents under the myopic best-response update rule. We have shown that based on the type of the population, the total number of cooperators in the long run either becomes fixed or fluctuates between two consecutive integers. The fluctuation is caused by a number of agents with a non-integer temper who switch between cooperation and defection. This is because when each such agent is playing cooperation and is activated to update her choice, she counts herself as one of the existing cooperators in the whole population. If instead, every agent excludes herself from the total number of cooperators when updating her choice and only focuses on the number of cooperators in the rest of the population, then there will be no fluctuation, and the state of the system \( x(t) \) will always converge to an equilibrium [42].

Our results on the Lyapunov-like function \( h(x) \) confine the asymptotic and more importantly transition behavior of the dynamics. According to the results, the agents may switch their strategies repeatedly for a while, but there are time instants after which some of the most cooperative agents, i.e., those with the highest tempers, will fix their choices to cooperation and some of the least cooperative agents will fix their choices to defection. The number of agents with fixed choices increases over time, although not necessarily monotonically. Hence, eventually all agents with tempers higher (lower) than that of the benchmark agents will cooperate (defect), resulting in the unique set of final states, regardless of the activation sequence of the agents. The findings, therefore, generalize those on homogeneous populations of snowdrift-game-playing individuals where a number equal to the temper of the individuals cooperate, and the rest defect in the long run (or they fluctuate between cooperation and defection if the temper is non-integer). It is of great interest to investigate whether such benchmark types of individuals also appear in structured populations.

We have used the convergence results to mathematically explain the level-off phenomenon, and have shown the existence of sustainable levels of cooperation where increasing the ratio of the less- or more-cooperative agents does not necessarily change the number of cooperators in the long run. This sheds light on how certain cooperation levels can be maintained in a population of selfish individuals. It can also explain why sometimes increasing the tendencies of some individuals in social groups does not lead to a higher level of cooperation. Moreover, the result indicates that if one wants to control the cooperation level of such populations by influencing the types of the agents, different choices of manipulated agents might lead to different outcomes. We are now working on formulating such control problems more formally and look for ultimately optimal control solutions by means of providing incentives to the individuals.

APPENDIX

A. Proof of Lemma 1

The update rule (2) dictates that agent \( i \) chooses the strategy that provides her the highest payoff against the population. Hence, according to the definition of \( u_i \), agent \( i \) chooses the strategy corresponding to the bigger entry of \( A_i s_C(t) \):

\[
A_i s_C(t) = \begin{bmatrix} R_i n_C(+1) + S_i (n - n_C(+1)) \\
T_i n_C(+1) + P_i (n - n_C(+1)) \end{bmatrix}.
\]

If agent \( i \) is a PD agent, from (3) it follows that \( T_i > R_i > P_i > S_i \). So the second entry is always bigger, leading to (6). On the other hand, if agent \( i \) is an SD \( j \) agent, using (45) to rewrite (2) gives:

\[
s_i(t+1) = \begin{cases} s^C & (R_i + P_i - T_i - S_i)n_C(t) + n_C(t) > n(P_i - S_i) \\
S_i(t) & (R_i + P_i - T_i - S_i)n_C(t) = n(P_i - S_i) \\
S^D & (R_i + P_i - T_i - S_i)n_C(t) < n(P_i - S_i) \end{cases}
\]

From (3) we know that for an SD \( j \) agent, it holds that

\[
T_i > R_i, \quad S_i > P_i \implies \begin{cases} R_i - T_i + P_i - S_i < 0 \\
S_i - P_i < 0 \end{cases}
\]

So

\[
s_i(t+1) = \begin{cases} s^C & n_C(t) < n_{SD} (\frac{n - S_i}{T_i - R_i + S_i - P_i}) \\
S_i(t) & n_C(t) = n_{SD} (\frac{n - S_i}{T_i - R_i + S_i - P_i}) \\
S^D & n_C(t) > n_{SD} (\frac{n - S_i}{T_i - R_i + S_i - P_i}) \end{cases},
\]

which is the same as (7) because of (4).

B. Proof of Lemma 3

First we prove (11). At \( x^* \), the number of cooperators in the population satisfies

\[
n_C(t_0) = \sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{kp}}^* \geq \lfloor n_{SD_{kp}}^* \rfloor.
\]
Similarly, at both \( x^* \) and \( x^*_+ \), it holds that
\[
n_C(t_0) \geq \sum_{j=1}^{k_p-1} n_{SD_j} + ([n_{SD_{dp}}^*] - \sum_{j=1}^{k_p-1} n_{SD_j}) = [n_{SD_{dp}}^*].
\]
So at \( x^* \), \( x^*_+ \) and \( x^*_+ \), it is true that
\[
n_C(t_0) \geq [n_{SD_{dp}}^*],
\]
which in view of (5) implies that
\[
n_C(t_0) > n_{SD_i}^* \quad \forall i > k_p.
\]
So from Lemma 1 we know that if an \( SD_i \), \( i > k_p \), agent is active at \( t_0 \), her strategy at \( t_0 + 1 \) remains to defect, which proves (11).

Now we prove (12). At \( x^* \),
\[
n_C(t_0) = \sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{dp}}^* - 1,
\]
while at both \( x^*_+ \) and \( x^*_+ \),
\[
n_C(t_0) \leq \sum_{j=1}^{k_p-1} n_{SD_j} + ([n_{SD_{dp}}^*] - \sum_{j=1}^{k_p-1} n_{SD_j}) = [n_{SD_{dp}}^*]
\]
which in view of (5) implies that
\[
n_C(t_0) < n_{SD_i}^* \quad \forall i < k_p.
\]
So from Lemma 1 we know that if an \( SD_i \), \( i < k_p \), agent is active at \( t_0 \), her strategy at \( t_0 + 1 \) remains to cooperate, which proves (12).

C. Proof of Lemma 8

When the population is clean-cut,
\[
n_C(t_0) = \sum_{j=1}^{k_p-1} n_{SD_j} \Rightarrow n_C(t_0) < n_{SD_{dp}}^* - 1.
\]
Otherwise,
\[
n_C(t_0) \in \{[n_{SD_{dp}}^*], [n_{SD_{dp}}^*] \} \Rightarrow n_C(t_0) \leq [n_{SD_{dp}}^*]
\]
which together imply that
\[
n_C(t_0) < n_{SD_{dp}}^* - 1.
\]

Hence, (38) holds for all unbiased populations.

D. Proof of Lemma 9

We prove by contradiction. Assume on the contrary that (9) can be violated. Then let \( t_1 \geq t_b + 1 \) be the first such violation time. So
\[
n_C(t_1) \geq n_{SD_{α(t_b)}}^* \Rightarrow n_C(t_1) \geq [n_{SD_{α(t_b)}}^*].
\]
However, from Lemma 8 we know that
\[
n_C(t_b) < n_{SD_{dp}}^* - 1,
\]
and since (24) is in force for all \( α \) including \( α = 0 \), in view of (5), we obtain
\[
n_C(t_b) < n_{SD_{α(t_b)}}^* + 1.
\]

By comparing \( n_C(t_b) \) and \( n_C(t_1) \) from (48) and (47) and knowing that \( n_C \) changes at most by one per time, we obtain
\[
\exists t' \in [t_b, t_1] : n_C(t') = [n_{SD_{α(t_b)}}^*] - 1. \tag{49}
\]
Let \( t_s \) be the greatest of all such possible \( t' \). Then in view of (47),
\[
n_C(t) \geq [n_{SD_{α(t_b)}}^*] \quad \forall t \in [t_s + 1, t_1]. \tag{50}
\]

Now we show that if an agent changes her strategy from \( D \) to \( C \) within \([t_s + 1, t_1]\), it must be an \( SD_{α(t_b)}+1 \) agent. Let \( t_c \) denote some time in \([t_s + 1, t_1]\). For an \( SD_j \), \( j > α(t_b) + 1 \), agent, it follows that
\[
(50) \Rightarrow n_C(t_c - 1) \geq [n_{SD_{α(t_b)}}^*] - 1 \tag{51}
\]
which, from Lemma 1, implies that the \( SD_j \) cannot change her choice to \( C \) at \( t_c \). On the other hand, for an agent \( SD_j \), \( j < α(t_b) + 1 \), we have
\[
j \leq α(t_b) \Rightarrow n_{SD_{Dj}}^*(t_b) = n_{SD_j}, \tag{52}
\]
and by the definition of \( t_1 \), since \( t_c - 1 \leq t_1 \),
\[
n_{SD_j}^*(t_c - 1) = n_{SD_{Dj}}^*(t_b), \tag{53}
\]

which together imply that
\[
n_{SD_j}^*(t_c - 1) = n_{SD_j} \quad \forall t_c \in [t_s + 1, t_1]. \tag{54}
\]

Therefore, all \( SD_j \) agents are already cooperators at \( t_c - 1 \) and hence do not change their choices to \( C \) at \( t_c \). So only an \( SD_{α(t_b)+1} \) agent can change her choice from \( D \) to \( C \) within \([t_s + 1, t_1]\). Now if \( t_1 > t_s + 1 \), from Lemma 1, (50) implies that no \( SD_j \) agent changes her choice to \( C \) during \([t_s + 2, t_1]\). Hence, no agent changes her choice from \( D \) to \( C \) during \([t_s + 2, t_1]\). So
\[
n_C(t) \leq n_C(t_s + 1) \quad \forall t \in [t_s + 1, t_1]. \tag{55}
\]

On the other hand, (55) also holds if \( t_1 = t_s + 1 \), verifying it in general. Now, (50) and the equality \( n_C(t_b) = [n_{SD_{α(t_b)}}^*] - 1 \) imply
\[
n_C(t_s + 1) = [n_{SD_{α(t_b)}}^*]. \tag{56}
\]

Then (50), (55) and (56) lead to
\[
n_C(t) = [n_{SD_{α(t_b)}}^*] \quad \forall t \in [t_s + 1, t_1]. \tag{57}
\]

So
\[
x(t_s + 1) = x(t_1)
\]
implicating that
\[
n_C(t_b) = n_C(t_1) \geq n_{SD_{α(t_b)}}^* + 1,
\]
which is a contradiction since \( t_s < t_1 \) and \( t_1 \) was defined as the first time after \( t_b \) at which \( n_C \) becomes non-less than \( n_{SD_{α(t_b)+1}} \).
Ming Cao is currently a professor of systems and control with the Engineering and Technology Institute (ENTEG) at the University of Groningen, the Netherlands, where he started as a tenure-track assistant professor in 2008. He received the Bachelor degree in 1999 and the Master degree in 2002 from Tsinghua University, Beijing, China, and the PhD degree in 2007 from Yale University, New Haven, CT, USA, all in electrical engineering. From September 2007 to August 2008, he was a postdoctoral research associate with the Department of Mechanical and Aerospace Engineering at Princeton University, Princeton, NJ, USA. He worked as a research intern during the summer of 2006 with the Mathematical Sciences Department at the IBM T. J. Watson Research Center, NY, USA. He is the 2017 and inaugural recipient of the Manfred Thoma medal from the International Federation of Automatic Control (IFAC) and the 2016 recipient of the European Control Award sponsored by the European Control Association (EUCA). He is an associate editor for IEEE Transactions on Automatic Control, IEEE Transactions on Circuits and Systems and Systems & Control Letters, and for the Conference Editorial Board of the IEEE Control Systems Society. He is also a member of the IFAC Technical Committee on Networked Systems. His main research interest is in autonomous agents and multi-agent systems, mobile sensor networks and complex networks.