Impact of deterministic and stochastic updates on network reciprocity in the prisoner's dilemma game

Jun Tanimoto*
Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Kasuga-koen, Kasuga-shi, Fukuoka 816-8580, Japan
(Received 5 April 2014; published 5 August 2014)

In 2 × 2 prisoner's dilemma games, network reciprocity is one mechanism for adding social viscosity, which leads to cooperative equilibrium. This study introduced an intriguing framework for the strategy update rule that allows any combination of a purely deterministic method, imitation max (IM), and a purely probabilistic one, pairwise Fermi (Fermi-PW). A series of simulations covering the whole range from IM to Fermi-PW reveals that, as a general tendency, the larger fractions of stochastic updating reduce network reciprocity, so long as the underlying lattice contains no noise in the degree of distribution. However, a small amount of stochastic flavor added to an otherwise perfectly deterministic update rule was actually found to enhance network reciprocity. This occurs because a subtle stochastic effect in the update rule improves the evolutionary trail in games having more stag-hunt-type dilemmas, although the same stochastic effect degenerates evolutionary trails in games having more chicken-type dilemmas. We explain these effects by dividing evolutionary trails into the *enduring* and *expanding* periods defined by Shigaki *et al.* [Phys. Rev. E **86**, 031141 (2012)].

DOI: 10.1103/PhysRevE.90.022105 PACS number(s): 02.50.Le

I. INTRODUCTION

Evolutionary games, such as prisoner's dilemma (PD), have been regarded as good metaphors to model a solution for the mysterious puzzle of why human beings and other animal species successfully evolve cooperation instead of egocentric defection within their societies. Many papers (for comprehensive reviews, refer to Refs. [1–3]) have discussed network reciprocity, which is one of the five fundamental mechanisms (and is believed to be the most important) that Nowak classified [4] for resolving the puzzle; it attempts to do so by adding "social viscosity." Network reciprocity continues to attract considerable attention because, although the central assumption of the model, i.e., playing with the neighbors on an underlying network and copying a strategy from them, is simple, it still seems plausible for explaining why cooperation survives in any real context.

For the past several years, researchers have been concerned with identifying additional model frameworks that would enhance network reciprocity to levels above those found in the baseline spatial PD (SPD) game. Nevertheless, any substantial and comprehensive understanding on why and how network reciprocity is brought about has still not been established. To understand what happens under the name of network reciprocity, dividing an evolutionary path, which starts from an initial random state and progresses to a final equilibrium state, into two periods seems persuasive [5–7]. Thus, in this study, we follow Shigaki *et al.* [5] and divide the path into an *enduring period* (END) and an *expanding period* (EXP). The END is the initial period in the dynamics in which the global cooperation fraction (P_c) decreases from its value at the initial state. Perhaps, the initial state has an equal number of cooperators and defectors randomly assigned on an underlying network. The term EXP refers to the period following END in which P_c increases.

The models of SPD games are constructed on an underlying network in which evolutionary dynamics progresses under a set of well-defined rules. In addition, each model incorporates an approved strategy for update rules; at one extreme, we have purely deterministic updates and at the other, purely stochastic updates. The use of different update strategies leads to difficulty and even ambiguity in trying to compare the results from the different models [8,9]. Stochastic update rules tend to realize further network reciprocity as a heterogeneous topology that adds noise with some cooperative enhancement over what happens on a homogeneous network; however, stochastic updates may mask what happens in deterministic updates. Therefore, in starting a discussion on what occurs during the process of network reciprocity, it seems a good idea to initially avoid any stochastic effects. For example, *imitation max* (IM) is a deterministic updating rule in which a focal player i imitates the strategy with the maximum payoff among all the strategies taken by the focal player and his or her immediate neighbors. At the other extreme, pairwise Fermi (Fermi-PW) is a representative stochastic rule in which a focal player i adopts the strategy of a randomly chosen player j with probability calculated by a Fermi function. In addition to Fermi-PW, there are several other methods classified as stochastic update rules, including pairwise linear (linear-PW) in which a linear function is presumed instead of the Fermi function from Fermi-PW and roulette selection (roulette) in which a focal player chooses from among all the strategies taken by the focal player and his or her immediate neighbors with a probability proportional to the payoff [or, strictly speaking, to the payoff difference with the minimum of the neighbors' payoffs as well as birth-death and death-birth processes (e.g., Ref. [10])]. Both Fermi-PW and linear-PW contain two stochastic layers: One is the process for selecting a pairwise partner for strategy adaptation among all neighbors, and the other is the process for deciding whether or not (keep her own strategy) to copy from the partner. Considering what happens in a real social system, we might insist that stochastic updating should be assumed in models because any real events might more likely be probabilistic rather

*Corresponding author: tanimoto@cm.kyushu-u.ac.jp

than deterministic. This might be true if we knew every aspect of network reciprocity. However, if not, it seems better to assume a deterministic procedure that simplifies the mechanism connecting events and that excludes any masking resulting from the stochastic noise.

In this paper, we introduce a holistic model of strategy updating including both IM and Fermi-PW as extreme cases; this model allows us to continuously vary updating rules from entirely deterministic to entirely stochastic. We find that a rule which is almost deterministic, but does contain subtle stochastic elements, shows better results in terms of emerging cooperation. The paper is organized as follows. Section II describes our model and the simulation procedure, Sec. III presents and discusses the results, and Sec. IV draws conclusions.

II. MODEL SETUP

At every time step, an agent on a network plays PD games with immediate neighbors and obtains payoffs from all games. As the underlying topology, we use the two-dimensional (2D) lattice graph of degree $k = 8$. The total number of agents is set to $N = 10^4$, which has been confirmed to be sufficiently large to yield simulation results that are insensitive to system size. After gaming, each agent synchronously updates his or her strategy.

A. Game description

In a PD game, a player receives a reward (R) for each mutual cooperation (C) and a punishment (P) for each mutual defection (D). If one player chooses C and the other chooses D, the latter obtains a temptation payoff (T), and the former is labeled a sucker (S). Without losing mathematical generality, we can define a PD game space by presuming $R = 1$ and $P = 0$ as follows:

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 1 & -D_r \\ 1 + D_g & 0 \end{pmatrix}, \quad (1)$$

where $D_g = T - R$ and $D_r = P - S$, i.e., simply set a chicken-type dilemma and stag-hunt- (SH-) type dilemma, respectively [11]. We limit the PD game class by assuming $0 \le D_g \le 1$ and $0 \le D_r \le 1$.

B. Agent's updating strategy

The strategy of an agent, C or D, is refreshed after every time step in the following way.

(1) The focal agent x selects a maximum payoff candidate y as his or her pairwise opponent from among n_{PW} neighbors who are randomly selected from all the neighbors. The number of neighbors in his or her neighbor set is k; thus, $n_{\mathrm{PW}} \le k$.

(2) Based on the Fermi-PW function, the focal agent x may or may not copy the strategy of the pairwise opponent y. This decision is based on the difference in payoff between x and y,

$$p_{\mathrm{copy}}^{x \leftarrow y} = 1 \Big/ \left[1 + \exp\left(\frac{\pi_x - \pi_y}{\kappa} \right) \right], \quad (2)$$

where π and k mean payoff and noise coefficients, respectively.

Obviously, if $n_{\mathrm{PW}} = k$ and $\kappa \to 0$, (2) recovers IM. However, if $n_{\mathrm{PW}} = 1$, (2) becomes the usual Fermi-PW. Thus, in the series of simulations, we control the two parameters n_{PW} and κ to obtain the mixtures of deterministic and stochastic strategies.

C. Simulation procedure

Each simulation was performed as follows. We set the initial cooperation fraction $P_{c,\mathrm{initial}}$. Initially, $N P_{c,\mathrm{initial}}$ cooperators and $N(1 - P_{c,\mathrm{initial}})$ defectors were randomly distributed among N agents allocated on different vertices of the network. Several simulation time steps, or generations, were run until the frequency of cooperation reached quasiequilibrium. If the cooperation frequency continued to fluctuate, we used the average frequency of cooperation over the last 250 generations of a 10 000-generation run. We varied the dilemma strength to cover PD: $0 \le D_g \le 1$ and $0 \le D_r \le 1$. The results reported below were drawn from 100 realizations, i.e., each ensemble average was formed from 100 independent simulations.

III. RESULTS AND DISCUSSION

Let us first confirm the holistic picture. Figure 1(a) shows the average cooperation fraction (P_c) over the entire region of the PD: $0 \le D_g \le 1$ and $0 \le D_r \le 1$ (hereafter, we use AllPD to represent the average P_c over the entire PD area). The figure shows AllPD in the 2D plane of the parameters κ vs n_{PW} for $P_{c,\mathrm{initial}} = 0.5$. Panels 1(b) and 1(c) give the information rates (bit) for each two-layer process in strategy updating (see Sec. II); the information rate is defined below. Again, the first process randomly picks n_{PW} neighbors from among k and selects the largest payoff agent from among those to make him or her the pairwise partner against the focal agent. Panel 1(b) shows the information rate for the first process that can successfully select the most appropriate partner (indicating the largest payoff agent, in general, however, selecting a max-payoff neighbor does not necessarily lead to emerging cooperation) compared with the situation in which the focal player can only select one of the k agents randomly. Thus, for $n_{\mathrm{PW}} = 1$, the information rate must be 0. In contrast, for $n_{\mathrm{PW}} = k = 8$ since the nominee pool always contains the partner that should be selected and the first process identifies him or her, the information rate is $\log_2(8) = 3$ bit. Here, let us confirm what information rate means in our model. That is, the information entropy of certain information indicating its advantage as compared with the situation with no information (inevitably leading to a random selection). Again, in the case of $n_{\mathrm{PW}} = 1$, there is no choice for the focal player in terms of her pairwise opponent. Whereas, in the case of $n_{\mathrm{PW}} = k$, the focal player can always find the real maximum payoff candidate. This is why the information rates of former and latter cases must be 0 and 3 bits, respectively. Panel 1(c) shows the information rate of the second process in the update strategy; the second process is the pairwise comparison. For the sake of comparison, we assumed two cases for the payoff difference $\Delta\pi$ between the focal agent (x) and the pairwise partner selected by the first process (y). As κ increases, the process becomes probabilistically random because $p_{\mathrm{copy}}^{x \leftarrow y} \to 0.5$. Thus, the information rate of this process asymptotically approaches zero. It is natural that the process has a larger information rate as $\Delta\pi$ becomes large. For the changes in $\Delta\pi$, the second process only varies in the range of [0,1] bits, which is less than the information rate of the first process. Note that, in terms of the information

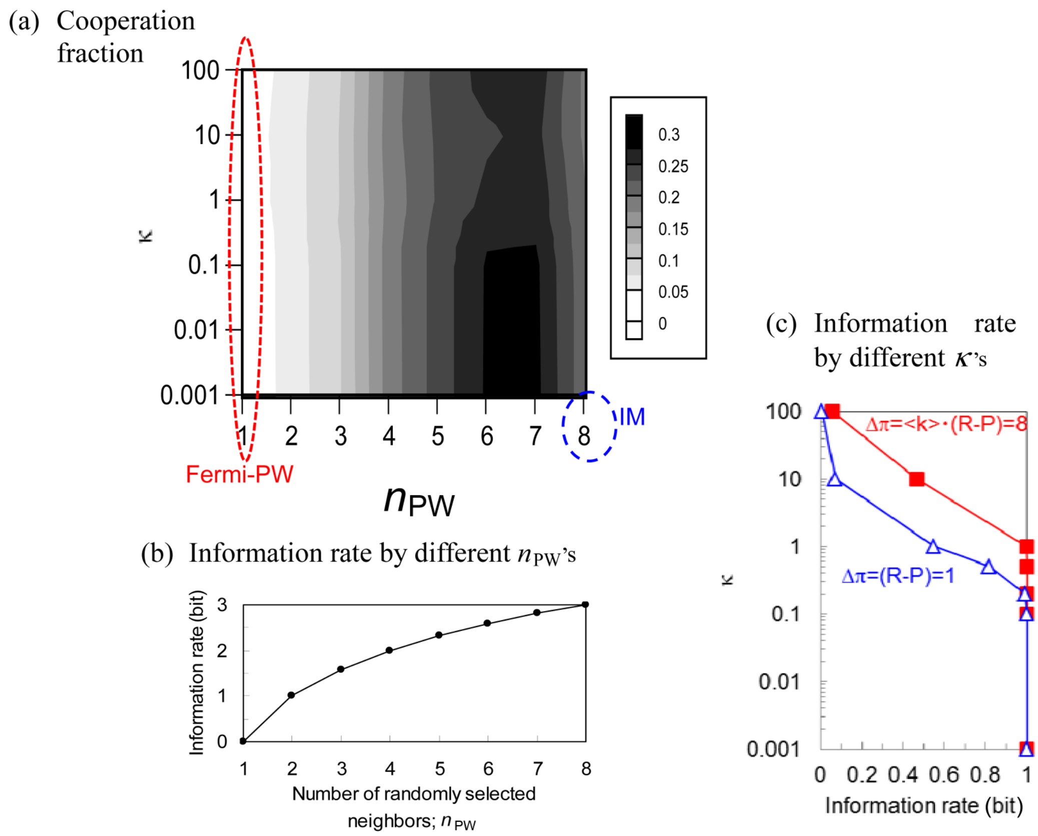

FIG. 1. (Color online) (a) Average cooperation fraction over the AllPD area on the $n_{\rm PW}$-κ plane for $P_{c,\rm initial} = 0.5$. Information rate as a function of (b) $n_{\rm PW}$ and (c) κ. In (c) the red line is for $\Delta\pi = \langle k\rangle(R-P) = 8$, and the blue line is for $\Delta\pi = 1(R-P) = 1$. In (a), the model recovers Fermi-PW if $n_{\rm PW} = 1$ and recovers IM if $n_{\rm PW} = 8$ and $\kappa \to 0$ simultaneously.

rate, the first process of selecting a final candidate for the second pairwise process is more meaningful than the second process of just choosing either themselves (the focal agent, i.e., keeping his or her own strategy in the process) or the final candidate so as to select the really appropriate agent from the set (i.e., one of all the neighbors plus the focal agent him- or herself). This point might somehow explain why a stochastically skewed selection of a pairwise opponent in the strategy adaptation process realizes more significant network reciprocity when a pairwise opponent, chosen as a reference for comparison, is selected, not randomly from all neighbors, but by a nonlinear proportional procedure applied to the payoff of each neighbor so as to make the high-payoff neighbor be selected as the pairwise opponent [12–14]. In fact, by carefully observing Fig. 1(a), the average cooperation fraction AllPD is noticed to be relatively sensitive to $n_{\rm PW}$ but insensitive to k. Thus, it appears that the player selected by the first process (such as an elimination race or a first-round match) serves a key role in enhancing network reciprocity. Contrariwise, the second process, which is the pairwise comparison (let us say, a final race), makes only a minor contribution. This is one of the things that is not correctly recognized so far.

According to the paper of Shigaki *et al.* [6], it seems a good idea to confirm what happens when a different $P_{c,\rm initial}$ is used. Figure 2 shows the average P_c AllPD analogous to Fig. 1(a) but for (a) $P_{c,\rm initial} = 0.25$ and (b) $P_{c,\rm initial} = 0.75$. Figures 1(a) and 2 show that sensitivity to the parameter $n_{\rm PW}$ is almost monotonic, but there is an obvious peak at approximately $n_{\rm PW} = 7$ irrespective of the value of $P_{c,\rm initial}$. More precisely, each peak in the three figures appears at approximately $n_{\rm PW} = 7$ and $\kappa = 0.1$.

Figure 3 shows P_c averaged over 100 realizations covering the full range of PD games $0 \le D_g \le 1$ and $0 \le D_r \le 1$ with fixed $\kappa = 0.1$. The upper, middle, and lower panels are for $P_{c,\rm initial} = 0.25$, 0.50, and 0.75, respectively. The left, center, and right panels are for $n_{\rm PW} = 1$, 7, and 8 (=k), respectively. Thus, panels 3(a), 3(d), and 3(g) are results using Fermi-PW with $\kappa = 0.1$, whereas panels 3(c), 3(f), and 3(i) are almost the same as the results using IM. Since Figs. 1 and 2 suggest that network reciprocity is rapidly degraded by the decreasing $n_{\rm PW}$, the results using Fermi-PW evidently look meager *vis-à-vis* those using IM. As long as we use a homogeneous network, such as the lattice, as the underlying network, IM leads to better network reciprocity than Fermi-PW. A deterministic update coupled with a homogeneous topology, such as the set of IM plus a lattice, enables geometrically less skewed expansions of C clusters in the EXP period that successfully survive the END period [5–7]. This improves the final equilibrium cooperation level over a stochastic update coupled with a homogeneous network, such as Fermi-PW plus lattice.

Comparing the results for $n_{\rm PW} = 7$ with those for $n_{\rm PW} = 8$, the cooperation is noted to expand its extent in the region close to the border with the SH game where D_r is relatively much larger than D_g. In contrast, cooperation is weakened in

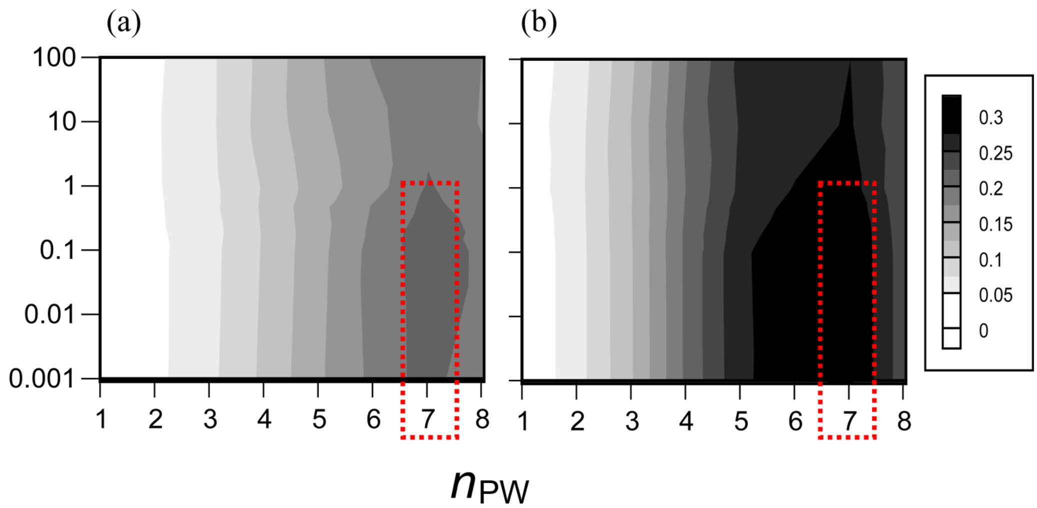

FIG. 2. (Color online) Average cooperation fraction of the AllPD area on the $n_{\rm PW}$-κ plane for (a) $P_{c,\rm initial} = 0.25$ and (b) $P_{c,\rm initial} = 0.75$.

the region close to the border with the chicken game where D_g is relatively much larger than D_r. Overall, the cases with $n_{\rm PW} = 7$ show peaks in terms of average cooperation fraction over the AllPD region as in Figs. 1 and 2. This finding might contribute to the comparison of entirely deterministic and slightly stochastic strategy update rules, although this peak effect, i.e., the further enhanced network reciprocity, is not so drastic compared with the entire deterministic case IM. In

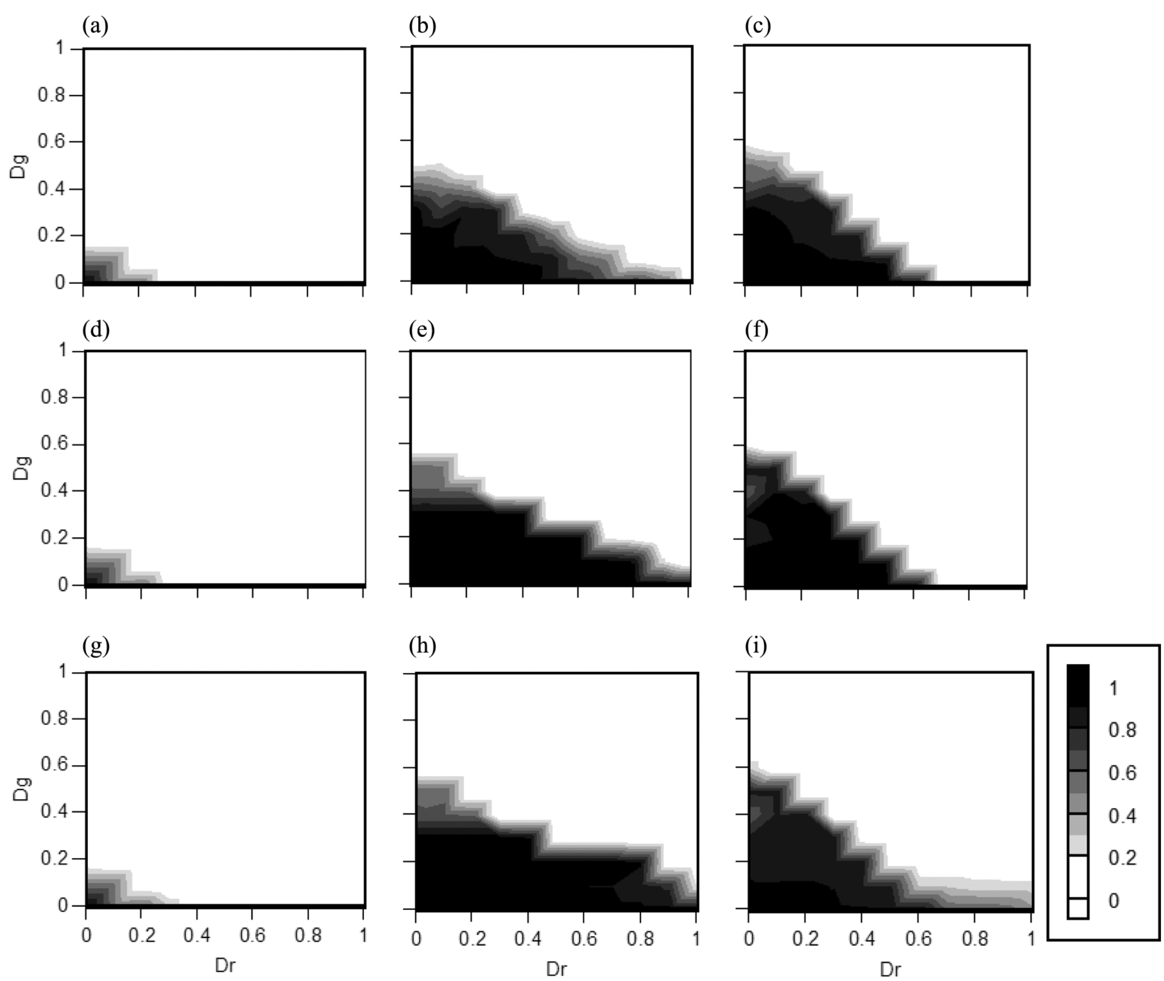

FIG. 3. Contour maps for cooperation fraction P_c over the entire range of PD games $0 \leqslant D_g \leqslant 1$ and $0 \leqslant D_r \leqslant 1$. Upper [(a)–(c)], middle [(d)–(f)], and lower [(g)–(i)] panels are for $P_{c,\rm initial} = 0.25, 0.50$, and 0.75, respectively. Left, center, and right panels are for $n_{\rm PW} = 1$, 7, and 8, respectively. All nine cases are for $\kappa = 0.1$. All results are ensemble averages over 100 realizations.

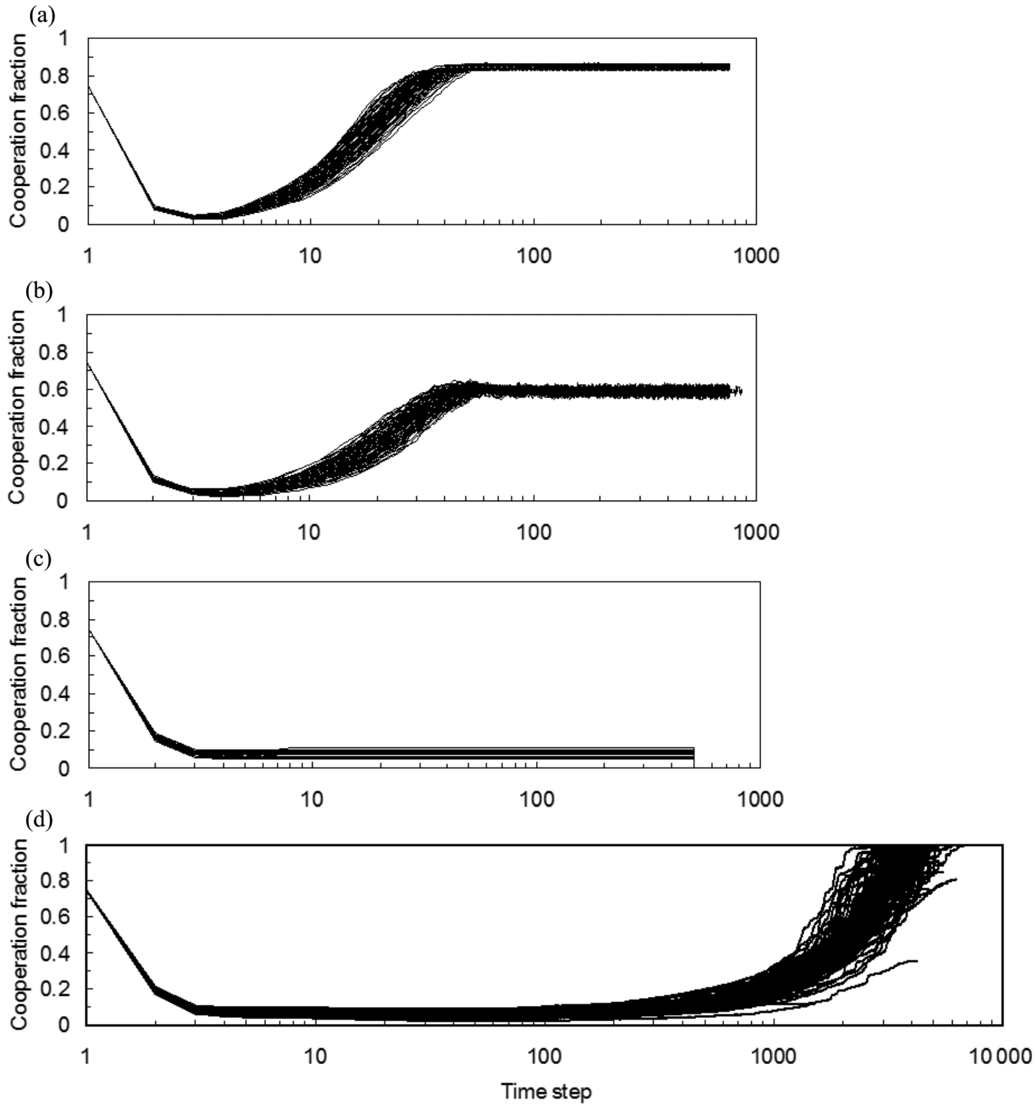

FIG. 4. Time evolutions for cooperation fraction P_c from 100 realizations with $\kappa = 0.1$ and $P_{c,\mathrm{initial}} = 0.75$. (a) $n_{\mathrm{PW}} = 8$ in a chicken-type dilemma with $D_g = 0.5$ and $D_r = 0.1$, (b) $n_{\mathrm{PW}} = 7$ in a chicken-type dilemma with $D_g = 0.5$ and $D_r = 0.1$, (c) $n_{\mathrm{PW}} = 8$ in a stag-hunt-type dilemma with $D_g = 0.2$ and $D_r = 0.8$, and (d) $n_{\mathrm{PW}} = 7$ in a stag-hunt-type dilemma with $D_g = 0.2$ and $D_r = 0.8$.

particular, so long as the homogeneous network is assumed, when a small amount of stochastic character is included in the updating, which is realized by setting $n_{\mathrm{PW}} = 7$ in our model, we obtain more network reciprocity than from the entirely deterministic update rule IM. Heretofore, the IM rule has been thought to be appropriate when coupled with a homogeneous network. Our observation can be paraphrased, such as this: Cooperation can be enhanced by adding a little noise to a system that is otherwise entirely stipulated by deterministic processes. This seems interesting from the statistical point of view. In the following, we try to explain how this can happen.

Figure 4 shows time evolutions of the cooperation fraction over 100 independent simulations with $n_{\mathrm{PW}} = 8$ and 7 and two representative dilemma structures: chicken-type dominant [panels 4(a) and 4(b)] with $D_g = 0.5$ and $D_r = 0.1$ and SH-type dominant [panels 4(c) and 4(d)] with $D_g = 0.2$ and $D_r = 0.8$. Panels 4(a) and 4(b) show that 100 realizations have almost the same equilibrium cooperation fraction when the chicken-type dilemma is dominant irrespective of n_{PW}. Panel 4(c) shows that the SH-type dilemma with $n_{\mathrm{PW}} = 8$ goes to a consistent but very meager level of cooperation. In contrast, Panel 4(d) with $n_{\mathrm{PW}} = 7$ realizes a fairly cooperative state, although each of the 100 realizations goes to a different equilibrium.

Let N_B be the number of agents who stand in the boundary between C and D, N_{CD} be the number of agents that change from C to D by the strategy copy process, and N_{DC} be the number of agents that change from D to C by the strategy copy process. Then, the fraction of agents who change from cooperation to defection is

$$F_{CD} = N_{CD}/N_B, \tag{3}$$

the fraction that changes from defection to cooperation is

$$F_{DC} = N_{DC}/N_B, \tag{4}$$

and the fraction that changes either way is

$$F = (N_{CD} + N_{DC})/N_B. \tag{5}$$

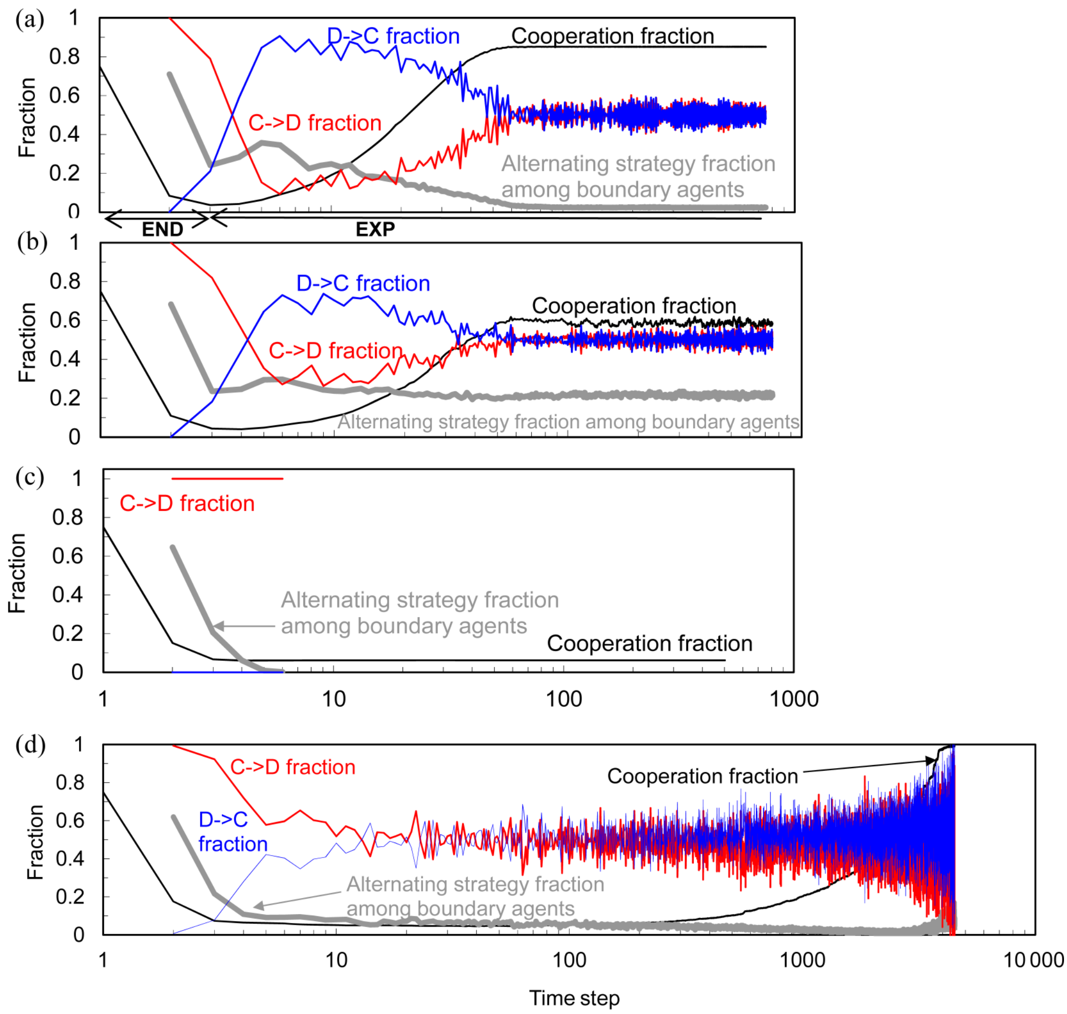

FIG. 5. (Color online) Representative time evolutions of cooperation fraction (P_c, black line), changing strategy fraction among boundary agents (F, gray bold line), shifting fraction from C to D (F_{CD}, red line), and shifting fraction from D to C (F_{DC}, blue line) with $\kappa = 0.1$ and $P_{c,\mathrm{initial}} = 0.75$. (a) $n_{\mathrm{PW}} = 8$ in a chicken-type dilemma with $D_g = 0.5$ and $D_r = 0.1$, (b) $n_{\mathrm{PW}} = 7$ in a chicken-type dilemma with $D_g = 0.5$ and $D_r = 0.1$, (c) $n_{\mathrm{PW}} = 8$ in a stag-hunt-type dilemma with $D_g = 0.2$ and $D_r = 0.8$, and (d) $n_{\mathrm{PW}} = 7$ in a stag-hunt-type dilemma with $D_g = 0.2$ and $D_r = 0.8$.

Figure 5 shows the representative time evolutions of these three fractions plus that for the cooperation fraction P_c in the same manner as in Fig. 4.

By comparing panels 5(c) and 5(d), we may be able to deduce how a little stochastic behavior [$n_{\mathrm{PW}} = 7$ in panel 5(d)] may foster more cooperation than a purely deterministic approach [$n_{\mathrm{PW}} = 8$ in panel 5(c)]. By assuming a slightly smaller number of neighbors [$n_{\mathrm{PW}} = 7$ in panel 5(d)] than the full number [$n_{\mathrm{PW}} = 8 = k$ in panel 5(c)], the defector who earns the maximum payoff among the neighbors is possibly excluded from the first process to select a pairwise opponent; this leads to a smaller probability that the focal cooperator copies defection from the max-payoff defector. This is meaningful during the END process. More importantly, this particular event only occurs when the SH-type dilemma dominates the chicken-type dilemma. Summing up, in this case, a little stochastic character in the process to select a pairwise opponent increases the cooperation level by improving the END period.

Likewise, by comparing panels 5(a) and 5(b), we can deduce how a little stochastic behavior can degrade cooperation for $n_{\mathrm{PW}} = 7$ relative to that for $n_{\mathrm{PW}} = 8$. By assuming a slightly smaller number of neighbors [$n_{\mathrm{PW}} = 7$ in panel 5(b)] than the full number [$n_{\mathrm{PW}} = 8 = k$ in panel 5(a)], the cooperator who earns the maximum payoff among the neighbors is possibly excluded from the first process to select a pairwise opponent; this leads to a larger probability that the focal cooperator copies defection from other neighbors rather than maintaining cooperation through the influence from the max-payoff cooperator. This is meaningful in the EXP process. In fact, the gap between F_{DC} and F_{CD} after moving from the END period into the EXP period for $n_{\mathrm{PW}} = 7$ is smaller than the gap for $n_{\mathrm{PW}} = 8$; this causes lower cooperative equilibrium for $n_{\mathrm{PW}} = 7$ than for $n_{\mathrm{PW}} = 8$. Summing up, when the chicken-type dilemma dominates the SH-type dilemma, a little stochastic character in the process to select a pairwise opponent reduces the cooperation level by degrading the EXP period.

In the end, the coexistence of these two contrasting effects, caused by injecting a little stochastic character into the strategy updating process, produces the peaks observed in Figs. 1(a), 2(a), and 2(b).

IV. CONCLUSIONS

We established a holistic framework for strategy updating that includes both IM and Fermi-PW as extreme cases; this framework allows us to continuously vary updating rules from entirely deterministic to entirely stochastic. By using this framework, we explored how any stochastic character in the updating process affects network reciprocity in PD games played on a homogeneous network.

We found that network reciprocity is improved overall when a little stochastic character is implemented so as to limit the number of potential candidates to be nominated as the pairwise opponent into the strategy updating process. This improvement occurs because injecting a small stochastic element, such as noise, into an entirely deterministic process realized by IM coupled with a lattice allows some sort of stochastic perturbation that increases the level of cooperation. With this insight, the effects can be decomposed into two contrasting elements; one is that a little stochastic flavor improves cooperation when a SH-type dilemma dominates, whereas a little stochastic flavor degrades cooperation when a chicken-type dilemma dominates.

Our results might contribute to a holistic view on how a small amount of stochastic character in a strategy updating process influences evolutionary trails that divide into END and EXP periods.

ACKNOWLEDGMENTS

This study was partially supported by a Grant-in-Aid for Scientific Research by JSPS, awarded to Professor Tanimoto (Grant No. 25560165), Tateishi Science & Technology Foundation. We would like to express our gratitude to these funding sources.

[1] G. Szabo and G. Fath, Phys. Rep. **446**, 97 (2007).
[2] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floría, and Y. Moreno, J. R. Soc., Interface **10**, 20120997 (2013).
[3] M. Perc and A. Szolnoki, BioSystems **99**, 109 (2010).
[4] M. A. Nowak, Science **314**, 1560 (2006).
[5] K. Shigaki, J. Tanimoto, Z. Wang, A. Hagishima, and N. Ikegaya, Phys. Rev. E **86**, 031141 (2012).
[6] K. Shigaki, Z. Wang, J. Tanimoto, and E. Fukuda, PLOS One **8**, e76942 (2013).
[7] Z. Wang, S. Kokubo, J. Tanimoto, E. Fukuda, and K. Shigaki, Phys. Rev. E **88**, 042145 (2013).
[8] A. Yamauchi, J. Tanimoto, and A. Hagishima, BioSystems **102**, 82 (2010).
[9] A. Yamauchi, J. Tanimoto, and A. Hagishima, BioSystems **103**, 85 (2011).
[10] M. A. Nowak, *Evolutionary Dynamics: Exploring the Equations of Life* (Belknap Press, Cambridge, MA, 2006).
[11] J. Tanimoto and H. Sagara, BioSystems **90**, 105 (2007).
[12] Z. Wang and M. Perc, Phys. Rev. E **82**, 021115 (2010).
[13] M. Perc and Z. Wang, PLOS One **5**, e15117 (2010).
[14] J. Tanimoto, M. Nakata, A. Hagishima, and N. Ikegaya, Physica A **391**, 680 (2012).